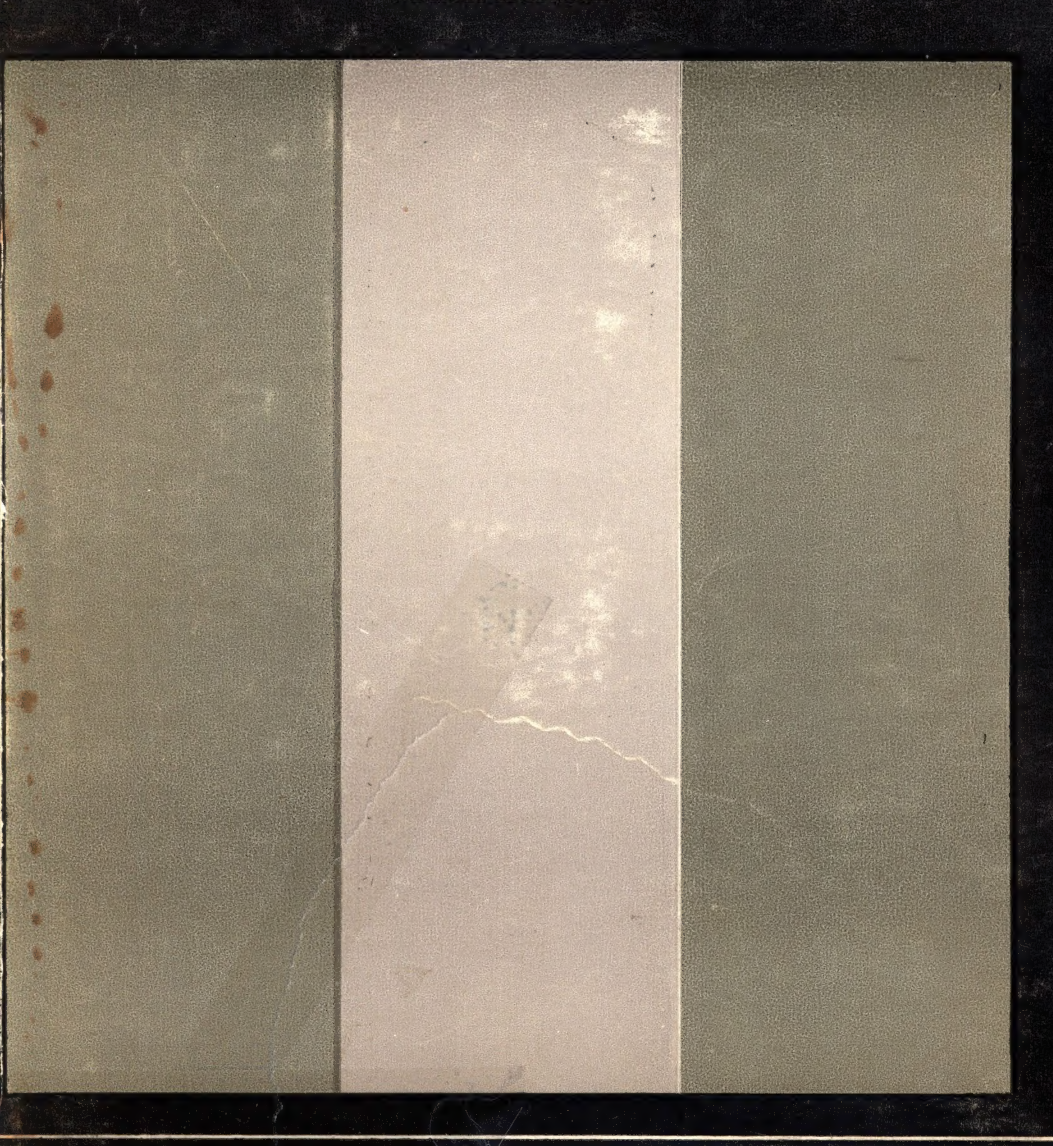
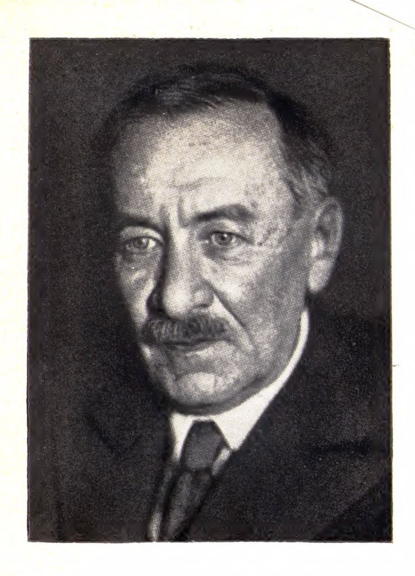
Академик

Л. И. МАНДЕЛЬШТАМ

ЛЕКЦИИ ПО ОПТИКЕ, ТЕОРИИ ОТНОСИТЕЛЬНОСТИ И КВАНТОВОЙ МЕХАНИКЕ





Леонид Исаакович Мандельштам (род. 5 мая 1879 г. в Одессе, умер 27 ноября 1944 г. в Москве) — выдающийся советский физик, широко известный своими трудами в области радиофизики, радиотехники, теории колебаний, оптики и молекулярной физики. Ученик Нобелевского лауреата Фердинанда Брауна (Страсбург).

С 1925 г. возглавил кафедру теоретической физики в Московском университете. В 1934 г. основал лабораторию колебаний в Физическом институте Академии наук СССР, носящую ныне его имя. С 1929 г. действительный член Академии наук СССР.

Л. И. Мандельштам создал одну из крупнейших в Советском Союзе физических школ, к первому поколению которой принадлежат, в частности, академики А. А. Андронов, Г. С. Ландсберг, М. А. Леонтович, Н. Д. Папалекси и И. Е. Тамм.

Многолетнее преподавание Л. И. Мандельштама в МГУ сыграло огромную роль в подъеме уровня физической науки в нашей стране и составляет столь же существенную власть его научного наследия, как и его научные результаты.

Академик
Л.И. МАНДЕЛЬШТАМ

ЛЕКЦИИ ПО ОПТИКЕ, ТЕОРИИ ОТНОСИТЕЛЬНОСТИ И КВАНТОВОЙ МЕХАНИКИ

АКАДЕМИЯ НАУК СССР

ОТДЕЛЕНИЕ ОБЩЕЙ ФИЗИКИ И АСТРОНОМИИ

Академик
Л. И. МАНДЕЛЬШТАМ

ЛЕКЦИИ ПО ОПТИКЕ, ТЕОРИИ ОТНОСИТЕЛЬНОСТИ И КВАНТОВОЙ МЕХАНИКЕ



ИЗДАТЕЛЬСТВО «НАУКА»
МОСКВА
1972

УДК 530.1

ПОД РЕДАКЦИЕЙ
члена-корреспондента АН СССР
С. М. РЫТОВА

2-3-4
249-71(I)

ПРЕДИСЛОВИЕ

Лекционные курсы и семинары, проведенные Леонидом Исааковичем Мандельштамом в Московском государственном университете в период с 1925 по 1944 г. (год его смерти), охватили чрезвычайно широкий круг проблем физики — теорию электромагнитного поля, электронную теорию, теорию колебаний, оптику, статистическую физику, специальную теорию относительности и квантовую механику. Значительную часть этих лекций и выступлений на семинарах удалось восстановить по записям слушателей, а в отдельных случаях — по стенограммам и по собственным конспектам и наброскам Л. И. Мандельштама. Этот материал составил содержание двух последних томов (IV и V) Полного собрания трудов Л. И. Мандельштама, опубликованного Издательством Академии наук СССР. В настоящее время все это издание, включая и оба указанные тома, вышедшие в 1955 г. (том IV) и 1950 г. (том V), сделалось библиографической редкостью.

Между тем лекции и семинары Л. И. Мандельштама ни в малейшей степени не утратили своего интереса и значения как по своему конкретному содержанию и глубине трактовки рассматриваемых физических вопросов, так и в качестве образца замечательного педагогического мастерства. В небольшом предисловии к тому V академик С. И. Вавилов писал: «... лекции и выступления на семинарах Л. И. Мандельштама представляют большую ценность и по существу затронутых вопросов, и по глубине и ясности их освещения. Общеизвестно, какое большое значение имели лекции и семинары Л. И. Мандельштама в Московском университете для развития теоретической физики. Эти лекции и семинары составляют очень важную часть научного наследства Л. И. Мандельштама наряду с его оригинальными научными трудами».

Здесь нет возможности подробно обрисовать все, что скрывается за этой сжатой характеристикой, рассказать о совершенно особой эмоциональной атмосфере, которую создавали лекции Л. И. Мандельштама, привлекавшие обширную аудиторию — не только студентов и аспирантов, но и многих зрелых научных работников. Представление обо всем этом можно почерпнуть из воспоминаний ряда учеников и сотрудников Л. И. Мандельштама [1]. Но необходимо хотя бы кратко пояснить, почему сейчас, спустя 25—45 лет со времени чтения этих лекций, Бюро Отделения общей физики и астрономии АН СССР нашло целесообразным вновь их переиздать.

Во-первых, как уже было сказано, они нисколько не устарели по содержанию. Они не уводят современного читателя в историю физики (хотя многие лекции и содержат исторический материал,

[1] См. «Изв. АН СССР», серия физич., 1945, 9, № 1-2, посвященный памяти Л. И. Мандельштама. Биография Л. И. Мандельштама, написанная академиками Н. Д. Папалекси и Г. С. Ландсбергом, помещена в I томе Полного собрания трудов. М., Изд-во АН СССР, 1948.

к тому же интересно освещенный), а показывают ему тот фундамент физической науки, на который она опирается и поныне.

Во-вторых, в подавляющем большинстве случаев в центре изложения находятся не фактические сведения сами по себе и не математический аппарат, а глубокий анализ содержания и развития основных физических теорий, конкретный анализ понятий, с которыми работает физик, т. е. все то, что составляет ядро физического мышления и что сплошь и рядом затрагивается лишь вскользь или вообще игнорируется в обычных учебниках.

Изложение специальной теории относительности, в котором подвергнуты внимательному разбору как исторические, так и логические ее предпосылки, в котором не обойдены, а, наоборот, тщательно проанализированы связанные с ней кажущиеся парадоксы и обычно возникающие психологические трудности,— это изложение является и сегодня одним из наилучших не только в отечественной, но, вероятно, и в мировой физической литературе. Лекции по основам квантовой механики в этом аспекте еще более значительны, так как здесь Л. И. Мандельштам развивает принадлежащую ему самому концепцию «косвенных измерений», которая, по общему признанию наших крупнейших «квантовиков», имеет глубочайшее значение для понимания того принципиально нового, что принесла квантовая механика.

И, наконец, сейчас, когда количество студентов, изучающих физику, и педагогов, преподающих ее в высшей школе, существенно возросло, лекции Л. И. Мандельштама, изданные увеличенным тиражом, принесут еще большую пользу, чем 15—20 лет тому назад, когда они были впервые опубликованы, причем тиражом всего в 3000 экземпляров. Несомненно, что для многих и многих из современной учащейся молодежи (а вероятно, и не только молодежи) чтение лекций Л. И. Мандельштама в ряде случаев явится откровением и заставит задуматься именно над тем, что достойно размышления, в отличие от ходовых «деловых» изложений, которые зачастую учат только считать, а не думать.

Лекции переизданы в двух книгах, воспроизводящих соответственно тома IV и V Полного собрания трудов Л. И. Мандельштама. Ссылки на тома I — III означают тома именно этого издания. Чтение одной книги не предполагает ознакомления с другой.

Конечно, за истекшие 20 лет многие направления физики далеко ушли вперед, но в задачу переиздания лекций не входили обширные комментарии, которые рисовали бы дальнейшее развитие затронутых в лекциях вопросов. Поэтому внесены лишь необходимые исправления и добавлены весьма немногочисленные новые примечания редактора, которые заключены в квадратные скобки.

С. М. Рытов

ЛЕКЦИИ ПО ИЗБРАННЫМ ВОПРОСАМ ОПТИКИ

(1932—1933 гг.)

ПЕРВАЯ ЛЕКЦИЯ

(19.X 1932 гг.)[1]

Роль парадоксов в развитии науки и при ее изучении. Содержание курса

Вопросы, которые мы будем рассматривать, выбраны по несколько необычному принципу. В истории развития науки большую роль играли противоречия взглядов — парадоксы. Два рода взглядов вступают в конфликт, и он вызывает дальнейшее движение науки вперед, дающее разрешение этого конфликта.

Примером этого является возникновение теории относительности. С одной стороны, все говорило за то, что скорость света относительно эфира не зависит от движения источника. К этому, например, приводили наблюдения над движением двойных звезд. Если бы это было не так, то наша электронная теория дисперсии и распространения света в веществе была бы невозможна из-за эффекта Допплера. Но тогда для наблюдателя, движущегося по отношению к эфиру со скоростью v, скорость света будет не c, а геометрическая сумма скорости c и скорости наблюдателя v, что несовместимо с результатом опыта Майкельсона. Этот конфликт был разрешен теорией относительности.

Конфликт, возникший в связи с теорией квант, тоже привел к развитию теории, к созданию новой квантовой механики. Приведем еще два примера.

История развития второго начала термодинамики. Карно (1824) стоял на точке зрения постоянства «количества теплоты» и руководился картиной, аналогичной падению воды с одного уровня на другой. Джоуль же экспериментально доказал возможность возникновения теплоты за счет совершения работы и таким образом заложил экспериментальные основы закона сохранения энергии. Эти два взгляда противоречат друг другу. Кельвин ясно видел это

---

[1] [Данный текст составлен по запискам С. М. Рытова и М. А. Леонтовича. Материал обработан М. А. Леонтовичем.]

противоречие. Создавая термодинамику, он долго боролся с этим противоречием. Сначала он пытался разрешить его, просто считая опыты Джоуля неверными. Конфликт разрешился в новой системе взглядов — в современной термодинамике.

Другой пример: H-теорема статистической механики. Больцман думал, что доказал эту теорему на основе механики. Однако Лошмидт указал, что достаточно в начальный момент обернуть знаки всех скоростей, чтобы процесс пошел в обратную сторону. Из результатов Пуанкаре, как это впервые указал Цермелло, следует, что рано или поздно система придет в первоначальное состояние. Далее Бертран указал Больцману, что при помощи теории вероятности из механики нельзя получить ничего нового (он иллюстрирует это тем примером, что, зная тоннаж и число труб судна, нельзя узнать, сколько лет капитану). Эти противоречия дали громадный толчок к развитию статистической механики.

Конфликты играли и более скромную роль, помогая уяснить картину явления. Наконец, есть еще более скромная область, где парадоксы также играют важную роль — это учебные вопросы. Есть две степени понимания. Первая, когда вы изучили какой-нибудь вопрос и как будто знаете все, что нужно, но вы еще не можете самостоятельно ответить на новый вопрос, относящийся к изучаемой области. И вторая степень понимания, когда появляется общая картина, ясное понимание всех связей. Такие вопросы, на которые нельзя ответить, пока этой второй степени понимания нет, мы называем парадоксами. Разбор подобных парадоксов очень полезен для достижения такого полного понимания.

Парадоксы есть во всех областях, но мы ограничимся некоторыми простыми вопросами из оптики. Конечно, цель наших занятий не решение ребусов, а разбор вопросов по существу. При этом вопросов чисто математического характера, вопросов, связанных со строгостью решения задач, мы не будем касаться.

Замечу, что в выборе вопросов для наших занятий есть известная старомодность,— мы будем разбирать главным образом вопросы классической физики. Однако многие из них имеют интерес и для квантовой физики, так как целый ряд задач волновой механики уясняется разбором вопросов классической оптики.

Охарактеризуем коротко те вопросы, которыми мы будем заниматься. Все вопросы оптики тесно между собой связаны, и нет, например, интерференции и дифракции как особых явлений, но все-таки некоторую классификацию оптических явлений, конечно, ввести можно.

Пусть имеется тонкая пластинка и отраженные ею лучи интерферируют. Разность фаз между лучами, отраженными от передней и задней ее плоскостей, будет при нормальном падении $4\pi\frac{d}{\lambda} - \pi$.

В зависимости от толщины пластинки d будет получаться усиление или ослабление отраженного луча. Амплитуда отраженного света будет заключена между $A + B$ и $A - B$, где A и B — амплитуды лучей, отраженных от задней и передней поверхностей. При этом не делается никаких предположений о величине A и B. Пусть нижняя поверхность идеально отражает, а сама пластинка ничего не поглощает (например, слой воды на ртути). Ясно, что свет должен полностью отразиться при любой толщине пластинки. Между тем предыдущее рассуждение приводит к тому, что амплитуда отраженного пластинкой света должна и в этом случае зависеть от ее толщины.

Заметим, что то обстоятельство, что при отражении должно происходить изменение фазы на π на одной из поверхностей раздела, вытекает из простых соображений. Если стремить толщину пластинки к нулю, то в пределе от нее не должно быть отражения. Это значит, что лучи, отраженные от ее двух поверхностей, имеют противоположные фазы. Но при бесконечно малой толщине пластинки разность фаз может получиться только из-за изменения фазы при отражении от поверхности.

Укажу еще на один парадокс из этой же области. Возьмем стеклянную пластинку, такую, что разность фаз лучей, отраженных ее передней и задней поверхностями, равна π. Через такую пластинку пройдет весь падающий на нее свет — она ничего не будет отражать. На пути луча можно поставить ряд таких пластинок, и так как результат не зависит от расстояния между ними, то их можно поставить вплотную. Мы получаем целое полупространство, заполненное прозрачным стеклом, от которого не получается никакого отражения. Между тем обычно говорят, что формулы Френеля *доказывают*, что при падении света на границу двух сред с разными показателями преломления всегда происходит отражение.

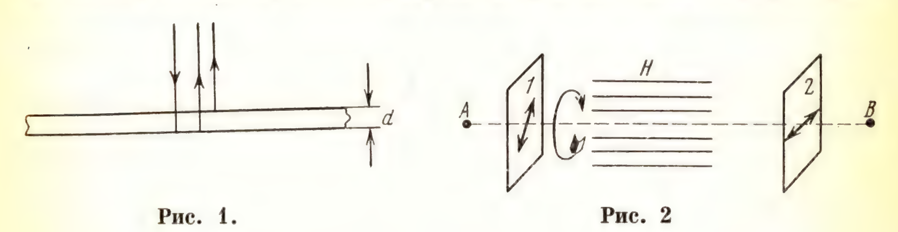


Рис. 1.   Рис. 2

Есть ряд более сложных вопросов, касающихся связи термодинамики с оптическими явлениями. Пусть имеются два излучающих осциллятора, две светящиеся точки. Если они излучают оба, то получится интерференционная картина. Энергия поля будет в

некоторых местах нуль, в других же — в четыре раза больше, чем в том случае, когда излучает один из них. Если подсчитать поток энергии через замкнутую поверхность, охватывающую оба источника, то он оказывается не равным сумме потоков энергии от каждого из них. Это ясно без вычислений в случае, когда расстояние между источниками значительно меньше половины длины волны. В этом случае, очевидно, ослабления поля нигде не будет. Фактически этот вопрос встал в беспроволочной телеграфии в случае двух антенн. Здесь как будто получается противоречие с законом сохранения энергии.

Другой курьезный вопрос возник у В. Вина. Он связан с явлением вращения плоскости поляризации света при распространении его в теле при наличии магнитного поля (явление Фарадея).

Магнитное вращение плоскости поляризации обладает следующим свойством, отличающим его от обычного вращения плоскости поляризации в оптически активных телах. При магнитном вращении знак угла поворота по отношению к определенному в пространстве направлению не зависит от направления луча. Посылая пучок света в одном направлении, отражая лучи и посылая их обратно, мы увеличиваем угол поворота вдвое по сравнению с его поворотом на пути в одном направлении. Поэтому можно осуществить такое устройство (рис. 2): николь *1* наклонен под углом 45° к горизонтальной плоскости, николь *2* — горизонтален, направление магнитного поля таково, что вращение плоскости поляризации происходит по стрелке на угол 45° при прохождении света от одного николя к другому (и притом в силу указанного свойства независимо от направления луча). Следовательно, точку A будет видно из B, а B из A не видно.

Поэтому Вин видел в этом устройстве способ осуществить perpetuum mobile второго рода, возможность осуществить передачу тепла от менее нагретого тела A к более нагретому B — без компенсации. Исходя из правильности второго начала термодинамики, он думал далее, что должны существовать какие-то еще неизвестные явления, за счет которых происходит компенсация этой передачи тепла к более нагретому телу. Однако Релей показал, что нет надобности искать какие-то неизвестные явления — такие компенсации имеются и вызваны хорошо известным явлением; они были упущены Вином только благодаря недостаточно подробному разбору этого устройства. Таким образом, этот perpetuum mobile второго рода работать не будет.

Перейдем к вопросам, связанным с модуляцией света. Если прерывать монохроматический свет или как-либо менять его амплитуду, то получающееся после этого колебание можно представить в виде

$$f(t)\cos \nu t,$$

Эту функцию всегда можно разложить на сумму незатухающих синусоидальных колебаний. Это математическая теорема. Вопросы и парадоксы возникают тогда, когда спрашивают, реальны ли эти слагаемые.

Опыты с модуляцией света делались при помощи конденсатора Керра. В этом случае получающиеся колебания таковы:

$$(1 + k \cos \omega t) \cos \nu t = \cos \nu t + \frac{k}{2} \cos (\nu + \omega) t + \frac{k}{2} \cos (\nu - \omega) t,$$

где ν — частота света, а ω — частота модуляции. Что значит «реальность отдельных членов» этой суммы? Это значит, что модулированный свет будет вести себя так же, как свет от трех монохроматических источников соответствующей частоты. Но вести себя можно по отношению к чему-либо, в данном случае — по отношению к исследующему этот свет прибору. Если взять дифракционную решетку (или другой спектральный прибор) достаточной разрешающей способности, то мы не сможем отличить модулированный свет и свет от трех таких источников. Следовательно, отдельные слагаемые разложения реальны.

Если это разложение реально, то возникает такой парадокс. Пусть источник начал светиться при $t = 0$. Световые колебания мы можем разложить на сумму монохроматических, т. е. синусоидальных, колебаний, длящихся от $t = -\infty$ до $t = +\infty$ (разложение в ряд или интеграл Фурье). Возьмем спектральный прибор большой разрешающей силы, например остро настроенный резонатор. Он выделит монохроматические колебания, и, следовательно, мы обнаружим световые колебания и до того момента, когда источник начал светиться. Это абсурд. Тогда такие разложения нереальны.

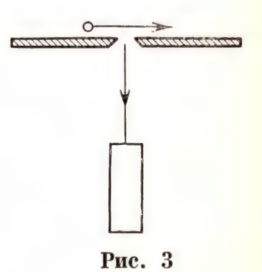


Рис. 3

На подобных же вещах основан известный парадокс с аномальной дисперсией и принципом относительности.

К этой же группе вопросов относится парадокс, указанный Эйнштейном. Пусть мимо щели пролетает светящаяся частица, излучающая монохроматический свет частоты ν, и мы наблюдаем сквозь щель посылаемый ею свет при помощи спектрального прибора (рис. 3). Если T — время пролета частицы мимо щели и если момент $t = 0$ соответствует положению частицы посредине щели, то световое поле за щелью будет

$$f(t) = 0, \quad |t| > \frac{T}{2},$$

$$f(t) = \sin \nu t, \quad |t| < \frac{T}{2}.$$

Свет за щелью будет не монохроматический, колебания в нем можно разложить в интеграл Фурье. Можно показать, что его спектр будет заключаться приблизительно в пределах частот $\nu - \frac{1}{T}$, $\nu + \frac{1}{T}$. Эйнштейн считал, что с точки зрения представления о световых квантах излучение света происходит мгновенно в виде монохроматического кванта и, следовательно, за щелью спектр должен быть узкий и указанного его размытия не должно быть. Этот опыт Эйнштейн считал своего рода experimentum crucis — волны или кванты. Опыт был сделан, и было обнаружено уширение линии.

Однако Бор указал, что и с точки зрения световых квантов уширение линии должно быть.

Дело в том, что благодаря дифракции будет свет, направленный в бок (не перпендикулярно плоскости щели), излученный частицей в направлении, не перпендикулярном направлению ее движения. Следовательно, будет эффект Допплера, а значит, будет и размытие линии. Подсчет показывает, что количественно отсюда получается ширина линии приблизительно такая же, как и из волнового представления.

Все это верно. Нехорошо только то, что и с волновой точки зрения есть и дифракция и эффект Допплера, а значит, и вызванное

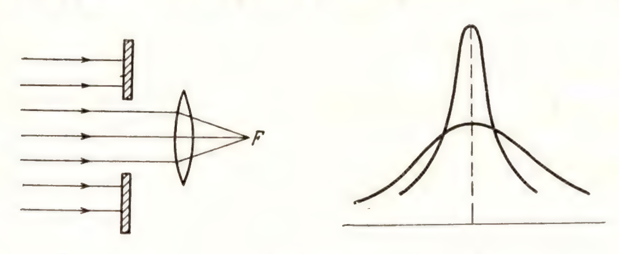


Рис 4.

ими уширение линии должно содержаться в классическом волновом рассмотрении явления. Но в классическом представлении явления, кроме того, есть вызванная прерыванием немонохроматичность, для которой нет соответствующего явления при квантовом представлении. Этот парадокс проистекает из учета одного и того же явления два раза — это часто встречающийся тип парадокса. Мы увидим, что прерывание или дифракция плюс допплер-эффект — это одно и то же.

«Световые биения» Риги — тоже недоразумение, связанное с этими же вопросами.

Перейдем теперь к группе вопросов из области теории дифракции. Эти вопросы играют роль и в оптике, и в акустике, и в волновой механике. Мы займемся ими в первую очередь.

Пусть имеется экран с отверстием, против которого помещена линза. Если удвоить площадь отверстия, то, по принципу Гюйгенса, амплитуда в фокусе должна удвоиться, и, следовательно, интенсивность учетверится. Однако ясно, что поток энергии через отверстие должен только удвоиться. Разрешение парадокса состоит в том, что понятие потока энергии через точку (фокус) лишено смысла. Нельзя отвлечься от дифракции. Благодаря дифракции свет будет не только в фокусе линзы, но и в известной области вокруг него. При увеличении площади отверстия вдвое амплитуда в фокусе увеличится вдвое, но область, в которой распределена энергия в фокальной плоскости, уменьшится как раз настолько, что полный поток энергии через фокальную плоскость лишь удвоится. Таким образом, в этой задаче, даже при сколь угодно большом отверстии, нельзя ограничиться применением геометрической оптики, нельзя пренебрегать дифракцией.

В заключение укажу еще на один квантовомеханический парадокс.

Пусть мы определяем координату электрона x, проходящего через отверстие в экране. Неточность в определении x-координаты, Δq, равна, очевидно, ширине отверстия. Из-за дифракции получится неточность в определении x-компоненты импульса p, равная

$$\Delta p = p \sin \varphi,$$

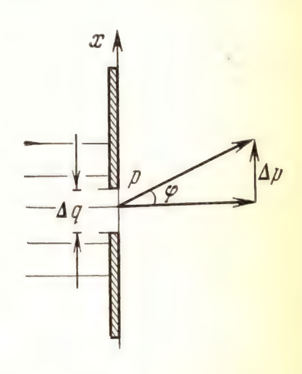


Рис. 5

где φ — угол дифракции. Таким образом, $|\Delta p| < p$. В то же время ширину щели и, значит, Δq можно сделать как угодно малой. Это, очевидно, противоречит соотношению неопределенностей

$$\Delta p \Delta q \geqslant h.$$

Мы увидим, что и это противоречие только кажущееся.

ВТОРАЯ ЛЕКЦИЯ

(1.XI 1932 г.)

Задачи теории дифракции. Метод их решений по Кирхгофу и границы его применимости. Когда имеет смысл говорить о потоке энергии через незамкнутую поверхность?

Сегодня мы будем заниматься вопросами дифракции, играющими большую роль в оптике и имеющими большое значение в акустике и в волновой механике.

Задача теории дифракции может быть формулирована приблизительно так. Имеется экран с отверстием, на него падает волна, требуется найти поле за экраном. Общие рецепты для решения таких дифракционных задач были даны еще Гюйгенсом и Френелем. Первый серьезный аппарат для решения этих задач был дан Кирхгофом, который получил соответствующие соотношения, исходя из волнового уравнения. В этом смысле Кирхгоф первый сделал попытку вывести принципы Гюйгенса и Френеля, позволяющие решать задачи дифракции из теории света.

Волновое уравнение

$$\nabla^2\varphi - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = 0 \qquad (1)$$

для монохроматического света $\varphi \sim e^{i\omega t}$ приобретает вид

$$\nabla^2\varphi + k^2\varphi = 0 \qquad (2)$$

$\left(k = \frac{\omega}{c} = \frac{2\pi}{\lambda}\right.$ — «циклическое» волновое число). Пользуясь теоремой Грина, можно показать, что значение решения φ волнового уравнения в некоторой точке P определяется через значения φ самого этого решения и его производной $\frac{\partial\varphi}{\partial n}$ по внешней нормали на замкнутой поверхности S, охватывающей область пространства, в которой находится точка P, и дается формулой

$$4\pi\varphi_P = \int\left\{\varphi\frac{\partial}{\partial n}\left(\frac{e^{-ikR}}{R}\right) - \frac{\partial\varphi}{\partial n}\,\frac{e^{-ikR}}{R}\right\}dS. \qquad (3)$$

Здесь R — расстояние точки P от элемента dS. Эту формулу Кирхгоф и применяет для решения задач дифракции, при этом он поступает следующим образом.

Формула Кирхгофа позволяет найти значение φ, если даны значения φ и $\frac{\partial\varphi}{\partial n}$ на замкнутой поверхности, например на замкнутом

экране с отверстием. Практически же наиболее интересен случай плоского бесконечного экрана. Можно этот случай свести к замкнутой поверхности, если замкнуть пространство за экраном с помощью бесконечно большой полусферы. Обыкновенно формулу (3) применяют к этому случаю, распространяя интеграл только по поверхности плоского экрана, между тем как он должен быть еще распространен по бесконечно большой полусфере. Обычно это не приводит к неправильному результату, так как интеграл по поверхности бесконечно большой полусферы, замыкающей экран, для большинства задач (хотя и не для всех) равен нулю. Примером, когда это не так, может служить случай, когда на экран падает волна не только слева, но и справа. Однако если считать, что волны на экран падают только с одной стороны, и искать решение, для которого за экраном есть только расходящаяся от него волна, то интеграл по бесконечной полусфере равен нулю. Только при этом необходимом дополнительном физическом требовании о *расходящихся* волнах получается единственное решение. Если этого требования не ставить, то можно получить много математически правильных решений, соответствующих разным физическим задачам.

Для доказательства исчезновения интеграла по поверхности бесконечно большой полусферы при условии расходящейся волны нужно учесть, что общий вид расходящейся волны такой:

$$\varphi = f(\theta,\ \psi)\,\frac{e^{-ikR}}{R}. \tag{4}$$

Если два решения φ_1 и φ_2 имеют этот вид при $R \to \infty$

$$\varphi_1 = f_1\frac{e^{-ikR}}{R}, \qquad \varphi_2 = f_2\frac{e^{-ikR}}{R},$$

то интеграл $\int\left\{\varphi_1\frac{\partial\varphi_2}{\partial n} - \varphi_2\frac{\partial\varphi_1}{\partial n}\right\}dS$, взятый по поверхности бесконечно большой полусферы, равен нулю. Этим и доказывается, что интеграл можно распространять в этом случае только по плоскости экрана.

Чтобы применить формулу (3), нужно знать φ и $\frac{\partial\varphi}{\partial n}$ на поверхности экрана и отверстия. Строго говоря, чтобы их знать, нужно уже знать решение задачи. Кирхгоф обходит эту трудность. Он дает рецепт, указывающий, как задавать φ и $\frac{\partial\varphi}{\partial n}$ на поверхности. Он говорит: для точек внутри отверстия возьмите те значения φ и $\frac{\partial\varphi}{\partial n}$, которые они имели в этом месте при той же падающей волне в отсутствие экрана, а вне отверстия, на экране, положите их

равными нулю. Например, для плоского экрана (расположенного в плоскости $z = 0$), на который нормально падает плоская волна $e^{i(\omega t - kz)}$, значения на отверстии надо брать равными

$$\varphi = e^{i\omega t}, \qquad \frac{\partial \varphi}{\partial n} = -ike^{i\omega t}. \tag{4'}$$

Если принять этот рецепт Кирхгофа, то решение задачи сводится к простому интегрированию по формуле (3). Это предположение, таким образом, дает возможность практически решить задачу, но, с другой стороны, оно влечет за собой целый ряд принципиальных трудностей.

Рассмотрим более подробно задачу о дифракции на одном отверстии в экране. Решение Кирхгофа относится к скалярной величине φ, но ведь в оптике световое поле определяется электрическим (**E**) и магнитным (**H**) векторами. Таким образом, для решения оптических задач, в сущности, нужно найти формулу, аналогичную формуле (3), но не для скаляра, удовлетворяющего волновому уравнению (2), а для векторов **E** и **H**, удовлетворяющих уравнениям Максвелла. Конечно, каждая из компонент **E** и **H** удовлетворяет волновому уравнению. Однако если просто применить формулу (3) к каждой компоненте **E** и **H**, то получается неправильный

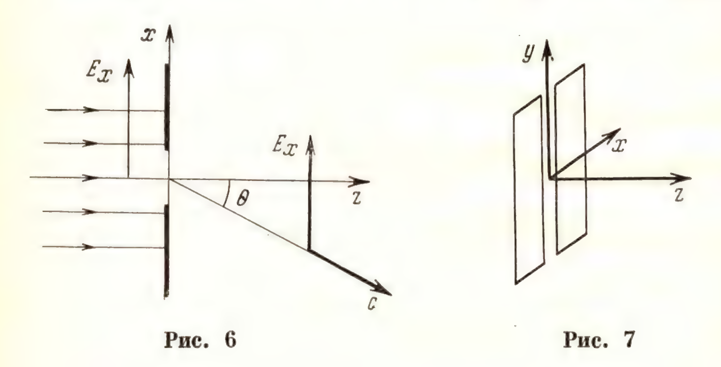


Рис. 6 Рис. 7

результат: во всем пространстве за экраном будут отличны от нуля только те компоненты поля, которые отличны от нуля в падающей волне. Это значит, что волны, идущие от отверстия, не будут волнами поперечными (рис. 6).

Здесь дело в том, что **E** и **H** удовлетворяют волновому уравнению, только если div **E** = 0 и div **H** = 0, т. е. только там, где нет зарядов. Решая, например, задачу о дифракции при нормальном падении волны на экран, расположенный при $z = 0$, по Кирхгофу, мы принимаем, что в плоскости $z = 0$ всюду $E_z = 0$, как и в па-

дающей волне. На краю экрана компоненты E_x и E_y испытывают в плоскости $z = 0$ скачки; это значит, что div E в этом месте не равна нулю, что соответствует наличию наведенных зарядов на краю экрана. Поэтому, решая волновые уравнения для E_x, E_y, E_z этим способом, мы и не получаем решения оптической задачи.

Однако есть один случай, когда это затруднение отпадает и который можно правильно решить, пользуясь скалярным волновым уравнением. Рассмотрим дифракцию на бесконечно длинной щели в экране, и пусть электрическое поле падающей волны параллельно щели (направленной вдоль оси y) и зависит только от x и z (рис. 7). В этом случае div **E**, очевидно, всюду равна нулю и задача (как и все плоские задачи) сводится к решению скалярного волнового уравнения. В этом случае при применении метода Кирхгофа нельзя отбрасывать интеграл по бесконечно удаленной поверхности полусферы, но теперь можно решать задачу с самого начала как задачу в двух измерениях. Волновое уравнение, которому удовлетворяет $E = E_y$, будет теперь

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial z^2} + k^2 E = 0. \tag{5}$$

При этом вместо формулы Кирхгофа (3) можно воспользоваться аналогичной формулой для волнового уравнения в двух измерениях, имеющей вид

$$2\pi\varphi_P = \int\left\{\varphi\,\frac{\partial Z(kr)}{\partial z} - Z(kr)\,\frac{\partial\varphi}{\partial z}\right\} ds, \tag{6}$$

где r — расстояние точки P от элемента замкнутого контура ds. Здесь $Z(\xi)$ с точностью до постоянного множителя — функция Ганкеля $Z(\xi) = \frac{i\pi}{2} H_0^{(2)}(\xi)$, встречающаяся во всех вопросах, связанных с распространением волн в двух измерениях. Для больших значений аргумента она равна

$$Z(\xi) \approx \sqrt{\frac{\pi}{2\xi}}\, e^{-i\left(\xi - \frac{\pi}{4}\right)}; \tag{7}$$

при малых ξ она ведет себя, как $\ln \xi$. В трех измерениях вспомогательная функция, входящая в формулу Кирхгофа, была e^{-ikR}/R и изображала расходящуюся шаровую волну, распространяющуюся от светящейся точки. В двухмерной задаче, когда поле не зависит от одной из трех координат, заменой светящейся точки является светящаяся линия. Функция Ганкеля $Z(kr)$ дает волну, распространяющуюся от светящейся линии.

Формулу (6) мы применим теперь к решению задачи о щели по методу Кирхгофа. При этом интеграл по бесконечно удаленной части контура интегрирования при тех же условиях, что и в трех измерениях, пропадает. И здесь нам придется воспользоваться основным допущением метода Кирхгофа — задать φ и $\frac{\partial\varphi}{\partial z}$ на щели такими же, как в ее отсутствие, и считать их нулем на экране. Это допущение является для большого числа задач единственным путем для их решения (так как строго решенных задач о дифракции не больше четырех-пяти), и его следствия хорошо оправдываются. Практически оно представляет собой гениальный шаг; однако именно в этом допущении кроется причина тех противоречий, к которым приводит иногда теория дифракции в форме Кирхгофа. Чтобы это показать, мы рассмотрим задачу о дифракции с энергетической точки зрения.

Для решения энергетических вопросов в электромагнитном поле нужно применить теорему Пойнтинга. Теорема Пойнтинга, которая доказывается, исходя из уравнений Максвелла, состоит в том, что прирост электромагнитной энергии W за единицу времени в объеме, ограниченном поверхностью σ, равен

$$\frac{\partial W}{\partial t} = -Q + \frac{c}{4\pi}\oint_{\sigma}[\mathbf{E}, \mathbf{H}]_n\, d\sigma, \tag{8}$$

где Q — приращение других видов энергии, например энергии тепловой. Для простоты мы положим $Q = 0$. Тогда, если ввести вектор Пойнтинга

$$\mathbf{S} = \frac{c}{4\pi}[\mathbf{E}, \mathbf{H}], \tag{9}$$

то

$$\frac{\partial W}{\partial t} = \oint_{\sigma} \mathbf{S}_n d\sigma. \tag{8'}$$

Эта теорема в виде (8′) — следствие уравнений Максвелла. Нужно обратить внимание на то, что при этом доказывается, что интеграл от нормальной компоненты вектора Пойнтинга по замкнутой поверхности дает прирост энергии внутри объема, ограниченного этой поверхностью. Эту теорему часто формулируют как утверждение, что $S_n d\sigma$ дает количество энергии, проходящее в единицу времени через площадку $d\sigma$. Конечно, соотношение (8′) этим будет удовлетворено, но при этом упускают, что обратное неверно. Такая локализация потока энергии отнюдь не вытекает из доказанного для замкнутой поверхности соотношения (8′), и поэтому такая локализация неоднозначна: (8′) будет выполнено

и в том случае, если вместо (9) принять

$$\bar{S} = \frac{c}{4\pi} [\mathbf{E}, \mathbf{H}] + \mathrm{rot}\, \mathbf{A}, \tag{9'}$$

где **A** — произвольный вектор.

Что же правильно? Какой поток энергии протекает через площадку $d\sigma$?

На этот вопрос можно ожидать одного из следующих трех ответов. Либо $\mathrm{rot}\, \mathbf{A} \neq 0$, либо $\mathrm{rot}\, \mathbf{A} = 0$, либо это безразлично, и вопрос лишен содержания. Правильно последнее, и поэтому все относящиеся сюда парадоксы не имеют смысла.

В учебниках часто говорят: примем, что $\mathrm{rot}\, \mathbf{A}$ прибавлять не нужно. Это неудачная формулировка, ибо при этом, очевидно, предполагается, что возможен опыт, который мог бы решить этот вопрос. Но для замкнутой поверхности такой опыт невозможен, так как из (8′) $\mathrm{rot}\, \mathbf{A}$ выпадает. В физических задачах никогда не имеет смысла говорить об изменении энергии внутри незамкнутой поверхности и о потоке через незамкнутую поверхность, так что вопрос о том, нужно ли к $\frac{c}{4\pi} [\mathbf{E}, \mathbf{H}]$ прибавить $\mathrm{rot}\, \mathbf{A}$, не может быть проверен на опыте [1].

Локализация потока энергии приводит к парадоксам потому, что из теоремы Пойнтинга выводят то, чего в ней не содержится, незаконно применяют соотношение, установленное для замкнутой поверхности, к отдельной площадке. Известный пример такого парадокса — комбинация непараллельных электрического и магнитного статических полей. Здесь $S = \frac{c}{4\pi} [\mathbf{E}, \mathbf{H}] \neq 0$, поток же вектора S через замкнутую поверхность, конечно, нуль. В связи с этим либо говорят, что к статическим полям рассуждения с вектором Пойнтинга неприменимы, либо же, что в статических полях электромагнитная энергия циркулирует по замкнутым кривым. Против первого утверждения можно возразить, что статическое поле — предельный случай переменного. Второе утверждение физически бессмысленно — энергия циркулирует, и это ни на чем не сказывается.

Укажем еще простой пример случая, когда локализация потока энергии на незамкнутой площадке часто приводит к недоразумениям. Представим себе передающую антенну (вытянутый проводник) с включенной в нее сторонней электродвижущей силой

---

[1] В гидродинамике в противоположность этому имеет смысл говорить о локализации потока жидкости, так как мы можем следить за движением определенной частицы жидкости.

(рис. 8). Если антенна—идеальный проводник, то вектор **E** перпендикулярен к поверхности антенны всюду, кроме участка, где действует сторонняя электродвижущая сила. Поэтому нормальная к поверхности антенны слагающая вектора Пойнтинга $S_n \sim [\mathbf{E}, \mathbf{H}]_n$ равна нулю на всех элементах поверхности антенны, кроме тех, где действует сторонняя электродвижущая сила. Полный поток вектора Пойнтинга через всю поверхность антенны дает правильный результат, он отличен от нуля — антенна излучает энергию. Но возьмем другой случай: в идеально проводящей антенне были возбуждены колебания и затем электродвижущая сила была выключена. Тогда во всех точках поверхности антенны $S_n = 0$ никакого потока энергии через поверхность антенны нет и как будто получается, что и излучения не будет.

Дело здесь просто. В объеме проводника антенны при выключенной электродвижущей силе нет никакого поля и, следовательно, нет и энергии, так что потока энергии через поверхность антенны в этом случае и не может быть. Но это не значит, что нет излучения. Излучение есть, и оно происходит благодаря тому, что энергия, находящаяся к моменту выключения электродвижущей силы в пространстве вблизи антенны, перемещается во все более и более далекие части пространства. Поток энергии через поверхность, окружающую антенну, но не совпадающую с ее собственной поверхностью, уже не будет нулем.

Недоразумения, вызванные локализацией вектора Пойнтинга, получаются и в оптике. Однако, в отличие от статического поля, в оптике имеет иногда смысл говорить о потоке энергии через незамкнутую площадку. Разбирая задачи дифракции, мы увидим, когда это имеет смысл. Мы увидим, что обычное толкование $S_n d\sigma = \frac{c}{4\pi} [\mathbf{E}, \mathbf{H}]_n d\sigma$ как потока энергии через площадку $d\sigma$ допустимо и имеет смысл тогда, когда размеры площадки $d\sigma$ велики по сравнению с длиной волны. В статическом случае длина волны равна бесконечности, и поэтому здесь локализация вектора Пойнтинга никогда не имеет смысла.

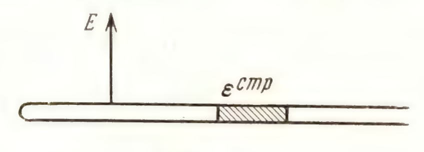


Рис. 8

ТРЕТЬЯ ЛЕКЦИЯ

(19.XII 1932 г.)

Дифракция на щели. Применение функции Грина в задачах дифракции

Рассмотрим дифракцию на щели плоской волны при нормальном ее падении и выясним энергетическую сторону дела в этом случае. Мы будем применять теорию Кирхгофа и воспользуемся формулой (6) второй лекции. Найдем поле на больших расстояниях от щели. В этом случае для $Z(kr)$ можно воспользоваться асимптотическим выражением

$$Z(kr) \approx \sqrt{\frac{\pi}{2kr}}\, e^{-i\left(kr - \frac{\pi}{4}\right)}.$$

Из чертежа видно, что

$$dz = -\frac{dr}{\cos\theta}$$

и, следовательно,

$$\frac{\partial Z}{dz} = -\frac{dZ}{dr}\cos\theta \approx \sqrt{\frac{\pi}{2kr}}\left(ik + \frac{1}{2r}\right)\cos\theta \cdot e^{-i\left(kr - \frac{\pi}{4}\right)}.$$

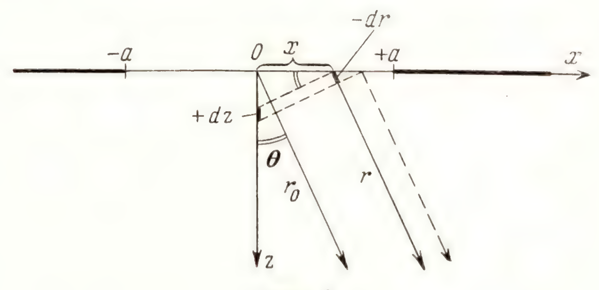


Рис. 9

Пренебрегая $1/2r$ по сравнению с $k = 2\pi/\lambda$, получаем

$$\frac{\partial Z}{\partial z} \approx i\sqrt{\frac{\pi k}{2r}}\cos\theta \cdot e^{-i\left(kr - \frac{\pi}{4}\right)}.$$

Подставив это выражение, (7) и (4′) в (6) (вторая лекция), получим

$$E = \frac{i}{2}\sqrt{\frac{k}{2\pi}}\int_{-a}^{+a} \frac{1+\cos\theta}{\sqrt{r}}\, e^{-i\left(kr - \frac{\pi}{4}\right)}\, dx.$$

Для очень больших расстояний $r = r_0 - x\sin\theta$, а под корнем r можно заменить на r_0; интегрируя, получим тогда

$$E = i\sqrt{\frac{k}{2\pi r_0}}\,(1+\cos\theta)\,e^{-i\left(kr_0-\frac{\pi}{4}\right)}\,\frac{\sin(ka\sin\theta)}{k\sin\theta}. \qquad (10)$$

Полученное выражение представляет собой цилиндрическую волну с амплитудой, зависящей от угла θ. Основная зависимость от угла θ дается типичным множителем $\frac{\sin(ka\sin\theta)}{k\sin\theta}$. В зависимости от угла амплитуда E имеет максимум при $\theta = 0$, равный $2a\sqrt{\frac{k}{2\pi r_0}}$. Острота этого максимума зависит от ka, т. е. от отношения ширины щели к длине волны. Если ka велико (ширина щели велика по сравнению с длиной волны), то поле за щелью сосредоточено в малых углах θ; в противоположном случае дифракционная картина размыта.

Рассмотрим баланс энергии при дифракции от щели. Применим теорему Пойнтинга к замкнутой поверхности — экран со щелью и бесконечно удаленная полуцилиндрическая поверхность. Мы вычислим потоки энергии через щель и через полуцилиндр, следуя Релею, который провел этот расчет и проверил, действительно ли здесь выполняется баланс энергии.

Нас будет интересовать среднее значение вектора Пойнтинга за период. Легко убедиться, что среднее значение вектора Пойнтинга за период равно[1]

$$S = \frac{c^2}{16i\omega\pi}\left([\mathbf{E}, \operatorname{rot}\mathbf{E}^*] - [\mathbf{E}^*, \operatorname{rot}\mathbf{E}]\right),$$

---

[1] Вычислим среднее значение, например, z-компоненты вектора Пойнтинга. Оно равно

$$\overline{S}_z = -\frac{c}{4\pi}\,\overline{\operatorname{Re}E_y\cdot\operatorname{Re}H_x}.$$

Здесь входят действительные части E_y и H_x, так как мы пользуемся комплексными величинами, и физический смысл имеют их действительные части. Так как

$$\operatorname{Re}E_y = \frac{1}{2}\,(E_y + E_y^*), \qquad \operatorname{Re}H_x = \frac{1}{2}\,(H_x + H_x^*)$$

и, кроме того, средние значения от E_yH_x и $E_y^*H_x^*$ равны нулю, а $E_yH_x^*$ и $E_y^*H_x$ не зависят от времени, то

$$\overline{S}_z = -\frac{c}{16\pi}\,(E_yH_x^* + E_y^*H_x).$$

В силу уравнений Максвелла

$$i\frac{\omega}{c}\,H_x = \frac{\partial E_y}{\partial z}, \qquad -i\frac{\omega}{c}\,H_x^* = \frac{\partial E_y^*}{\partial z},$$

где ω — частота, а звездочкой обозначены комплексно сопряженные величины. Энергия, проходящая через щель, равна (множитель $c^2/16\,\omega\pi$ мы будем всюду отбрасывать)

$$J_0 = \frac{1}{i}\int_{-a}^{+a}\left(E\,\frac{\partial E^*}{\partial z} - E^*\,\frac{\partial E}{\partial z}\right)dx.$$

Так как на щели $E = 1$, $\frac{\partial E}{\partial z} = ik$, то мы получаем

$$J_0 = 4\,ka.$$

На основании теоремы Пойнтинга мы можем утверждать, что такой же поток энергии протекает и через бесконечно большую полуцилиндрическую поверхность. Вычислим, однако, поток через эту поверхность непосредственно.

Принимая во внимание, что на этой поверхности при $r_0 \gg \lambda$

$$\frac{\partial}{\partial n} = \frac{\partial}{\partial r_0}, \qquad \frac{\partial E}{\partial r_0} \approx k\sqrt{\frac{k}{2\pi r_0}}\,(1+\cos\theta)\,e^{-i\left(kr_0 - \frac{\pi}{4}\right)}\,\frac{\sin(ka\sin\theta)}{k\sin\theta},$$

получим для потока выражение

$$J = \frac{r_0}{i}\int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}}\left(E\frac{\partial E^*}{\partial r_0} - E^*\,\frac{\partial E}{\partial r_0}\right)d\theta = \frac{1}{\pi}\int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}}(1+\cos\theta)^2\left(\frac{\sin(ka\sin\theta)}{\sin\theta}\right)^2 d\theta.$$

Если щель очень велика по сравнению с длиной волны ($ka \gg 1$), то, в силу наличия очень острого максимума амплитуды поля при малых углах дифракции, в интеграле существенна только та его часть, которая относится к малым углам. Поэтому под интег-

---

так что

$$\overline{S}_z = \frac{c^2}{16i\omega\pi}\left(E_x\,\frac{\partial E_y^*}{\partial z} - E_y^*\,\frac{\partial E_y}{\partial z}\right).$$

Точно так же получим

$$\overline{S}_x = \frac{c^2}{16i\omega\pi}\left(E_y\,\frac{\partial E_y^*}{\partial x} - E_y^*\,\frac{\partial E_y}{\partial x}\right).$$

Легко видеть, что это компоненты векторного выражения, приведенного в тексте, в случае нашей двухмерной задачи ($E_y = E$).

ралом можно заменить $\sin\theta$ на θ, а $\cos\theta$ на 1. Сделав это, получим

$$J = \frac{4k^2a^2}{\pi} \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \left(\frac{\sin ka\theta}{ka\theta}\right)^2 d\theta = \frac{4ka}{\pi} \int_{-\frac{\pi}{2}ka}^{+\frac{\pi}{2}ka} \left(\frac{\sin\xi}{\xi}\right)^2 d\xi.$$

Пределы по ξ можно теперь в нашем случае приближенно заменить на $-\infty$ и на $+\infty$. Тогда, учитывая, что

$$\int_{-\infty}^{+\infty} \left(\frac{\sin\xi}{\xi}\right)^2 d\xi = \pi,$$

получим

$$J = 4ka.$$

Таким образом, все обстоит хорошо, баланс энергии соблюдается, теорема Пойнтинга справедлива.

Однако полученное совпадение даже слишком хорошо. Действительно, при расчете были сделаны пренебрежения и, следовательно, должно было получиться лишь приближенное совпадение J с J_0.

Рассмотрим обратный случай — щель, очень узкую по сравнению с длиной волны ($ka \ll 1$). В этом случае

$$\frac{\sin(ka\sin\theta)}{ka\sin\theta} = 1$$

и

$$J = \frac{k^2a^2}{\pi} \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} (1+\cos\theta)^2\, d\theta = \frac{3}{2} k^2a^2 \neq J_0.$$

Это грубое нарушение теоремы Пойнтинга, баланса энергии нет. В чем здесь дело?

Иногда на этот вопрос вам ответят так: неправильны те значения E и $\frac{\partial E}{\partial z}$, которые вы взяли на отверстии равными значениям при отсутствии экрана, а на экране — равными нулю. Говорят, что энергетическая проверка для широкой по сравнению с длиной волны щели показывает, что в данном случае это предположение Кирхгофа относительно E и $\frac{\partial E}{\partial z}$ на щели правильно,

Этот ответ неудовлетворителен. Нужно выяснить, что значит неправильное задание E и $\frac{\partial E}{\partial z}$ на экране и отверстии. Нужно охарактеризовать физические свойства экрана. Будем, например, считать экран идеально проводящим. Для решения задачи нужно найти поле E такое, которое вместе со своими первыми производными было бы всюду вне экрана (в том числе и на отверстии) непрерывно, на экране обращалось бы в нуль, а за экраном представляло бы собой расходящиеся волны. Решение этой задачи дает поле не такое, какое получается при решении по методу Кирхгофа.

Однако это еще не значит, что значения E и $\frac{\partial E}{\partial z}$, которые берет Кирхгоф на экране и в отверстии, неправильны — они могут соответствовать экрану с другими физическими свойствами, не идеально проводящему экрану. Может быть, можно придумать такой экран, на котором E и $\frac{\partial E}{\partial z}$ имели бы эти значения. Возможно ли это на самом деле, совместимо ли это с тем, что E удовлетворяет волновому уравнению? Это вопрос чисто математический.

Решение вопроса состоит в том, что это невозможно, что предположение Кирхгофа математически противоречиво. Кирхгофом было доказано только, что если E есть решение волнового уравнения, то между значениями E и $\frac{\partial E}{\partial n}$ на замкнутой поверхности и внутри нее имеет место соотношение (3) [или в двух измерениях — (6)]. Но он не доказал, что при произвольно заданных на поверхности функциях E и $\frac{\partial E}{\partial n}$ левая часть выражения (3) [или (6)] будет таким решением волнового уравнения, значения которого на поверхности будут равны этой функции E, а значения его производной — $\frac{\partial E}{\partial n}$.

Вообще говоря, это не имеет места. При решении волнового уравнения $\nabla^2 E + k^2 E = 0$ нельзя произвольно задавать на замкнутой поверхности E и $\frac{\partial E}{\partial n}$, так как решение волнового уравнения, вообще говоря, определяется заданием на этой поверхности одного только $\frac{\partial E}{\partial n}$ (или одного только E). Задание E (или $\frac{\partial E}{\partial n}$) на плоскости $z = 0$ и условие уходящих волн на бесконечно большой полусфере определяют решение волнового уравнения однозначно. Значит, предположение Кирхгофа, с одной стороны, не всегда может быть оправдано физически, а с другой стороны,— математически противоречиво. Кирхгоф знал

это, но считал, что с достаточным приближением его допущение все же справедливо.

Теорию Кирхгофа можно избавить от этой математической неприятности. Так как решение волнового уравнения определяется заданием значений одного E на плоскости $z = 0$, то можно найти представление решения только через эти значения. Для этого вместо $Z\,(kr)$ возьмем другую функцию

$$\Gamma = Z\,(kr) - Z',$$

где Z' — всюду непрерывное внутри поверхности решение волнового уравнения, представляющее собой на бесконечном полуцилиндре уходящие волны. Соотношение (6) останется тогда справедливо, если в нем заменить Z на Γ, так как для непрерывной функции Z': $\oint \left(E\frac{\partial Z'}{\partial z} - Z'\frac{\partial E}{\partial z}\right) ds = 0$. Можно положить

$$Z' = Z\,(kr'),$$

где r' — расстояние от зеркального изображения P' точки P в плоскости $z = 0$ (рис. 10). Тогда на плоскости $z = 0$, $\Gamma = 0$ и $\frac{\partial \Gamma}{\partial z} = 2\frac{\partial Z}{\partial z}$, а интеграл по бесконечному полуцилиндру по-прежнему пропадает. Поэтому (для двухмерного случая) получаем равенство

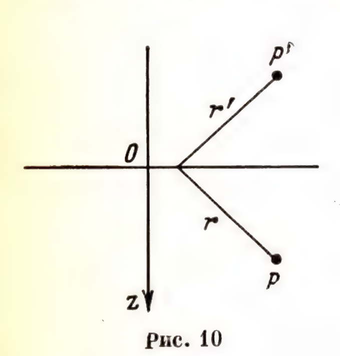


Рис. 10

$$\pi E_P = \int E\frac{\partial Z}{\partial z}\,dx. \tag{11}$$

Значение $\frac{\partial E}{\partial z}$ на поверхности выпало. Эта формула Грина дает возможность решать задачу математически корректно.

Однако и при применении вместо (6) этой формулы (11) остается вопрос о том, как задавать значение E на плоскости экрана с отверстием. Поэтому формула Грина, в сущности, мало подвинула нас к решению вопроса.

Поскольку задание E на плоскости $z = 0$ и условие уходящих волн на бесконечности определяют решение однозначно, постольку, если решение, удовлетворяющее этим условиям, будет каким-нибудь образом найдено, будет то решение, которое нужно. Поэтому можно получить это решение и не применяя формулы (11) и пользуясь, например, методом Релея (представляя решение в зависимости от x в виде ряда или интеграла Фурье).

Замечу, что я вижу в теории дифракции еще одну неприятность. Мы получили для щели решение (10), годное для очень больших расстояний от нее, для таких, что $r \gg \frac{a^2}{\lambda}$. Это так называемый случай Фраунгофера. Это решение дает нам распределение интенсивности на очень большом расстоянии от отверстия: если там поставить экран, то на нем будет широкое светлое пятно, распределение интенсивности в котором мы выяснили. Обычно для получения фраунгоферовой дифракции рекомендуют пользоваться трубой (телескопом), и говорят, что при этом получится то же самое, что и без трубы на очень большом расстоянии. Однако из изложенного это без дальнейшего еще не вытекает. Единственный случай, когда наблюдают дифракцию без трубы — это случай дифракции рентгеновских лучей от кристаллов. Однако в этом случае условие $r \gg \frac{a^2}{\lambda}$ не выполнено, так как здесь $\frac{a^2}{\lambda} \sim 10^7$ *см*. В чем здесь дело, будет объяснено дальше.

ЧЕТВЕРТАЯ ЛЕКЦИЯ

(25.XII 1932 г.)

Строгая постановка задачи о дифракции. О невозможности абсолютно черных экранов. Дифракция на очень малых телах. Рассеяние энергии резонатором и электроном

Математическую задачу, соответствующую физической задаче о дифракции плоской волны, падающей нормально на щель в идеально проводящем бесконечно тонком экране, можно формулировать так.

Требуется найти решение волнового уравнения

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial z^2} + k^2 E = 0,$$

удовлетворяющее следующим условиям.

При $z \to \infty$ (далеко за экраном)

$$E = f(\theta)\,\frac{e^{-ikr}}{\sqrt{r}},$$

т. е. поле должно представлять собой волну, расходящуюся от щели (r здесь — расстояние от щели).

При $z \to -\infty$

$$E = e^{-ikz} - e^{ikz},$$

т. е. на больших расстояниях от щели со стороны падения света поле должно стремиться к сумме падающей и отраженной от экрана волн.

На поверхности идеально проводящего экрана всюду $E = 0$, и, наконец, всюду вне экрана, включая и отверстие, E и его производные должны быть непрерывны.

Формулированная таким образом задача очень сложна. Идея Кирхгофа, развившего свой метод решения дифракционных задач, состояла в том, чтобы обойти трудности ее решения. Мы видели, что, вводя «функцию Грина» и применяя формулу (11), мы устраняем математические недостатки метода Кирхгофа. Однако необходимость заранее знать значения поля E в точках отверстия остается. Подробный анализ показывает, что предположение Кирхгофа для задания этих значений E с достаточным приближением справедливо, только если размеры отверстия велики по сравнению с длиной волны. Таким образом, в этой форме метод Кирхгофа может быть оправдан для больших отверстий.

Вообще для больших отверстий в экране и электрическое поле $\mathbf{E}$ и магнитное поле $\mathbf{H}$ будут такие же, как без экрана. Значит, в этом случае и вектор Пойнтинга в отверстии будет такой же, если бы экрана не было.

Как мы знаем, теорема Пойнтинга относится к замкнутой поверхности. Применим ее к замкнутой поверхности, образованной экраном с отверстием и полусферой бесконечно большого радиуса. Будем предполагать, что среда обладает некоторым поглощением (хотя бы и очень малым). Тогда интеграл по полусфере будет равен нулю, так что останется только интеграл по экрану и отверстию. При большом по сравнению с длиной волны отверстии правильно предположение Кирхгофа, поэтому поток через отверстие равен потоку через ту же площадку при отсутствии экрана.

Теперь мы можем указать те случаи, когда имеет смысл говорить о потоке энергии через незамкнутую пощадку. Это имеет смысл, если *размеры площадки велики по сравнению с длиной волны.*

Мы можем с помощью экрана с большим по сравнению с длиной волны отверстием выделить и измерить энергию, проходящую через площадку, совпадающую с отверстием, и утверждать, что в падающей волне и при отсутствии экрана через эту площадку будет проходить та же энергия.

Я избегал говорить о теле с абсолютно черной поглощающей поверхностью и провел рассуждение с поглощающей средой, так как абсолютно черная поверхность — функция, не совместимая с уравнениями Максвелла. Не всякие граничные условия совместимы с дифференциальными уравнениями. По Максвеллу, всякое тело характеризуется диэлектрической постоянной ε и проводимостью σ. Чтобы получить большое поглощение, нужно брать тело с большой проводимостью, но при увеличении проводимости увеличивается и отражение от поверхности тела и при $\sigma = \infty$ отражение делается полным. Если взять пластинку, у которой проводимость σ, а значит, и коэффициент поглощения можно менять, например кювету с раствором краски, и постепенно увеличивать ее коэффициент поглощения, то при определенном значении σ количество тепла, выделяющееся в пластинке за счет поглощения ею света, будет максимально. Действительно, и при $\sigma = 0$, когда весь свет проходит через пластинку, и при $\sigma = \infty$, когда весь свет отражается от нее, это тепло будет нулем. Впрочем, на практике поверхности, покрытые сажей, часто достаточно хорошо воспроизводят абсолютно черную поверхность.

Строгое решение дифракционных задач, вообще говоря, очень сложно. Проще других случай, когда рассеивающий объект очень мал по сравнению с длиной волны. В этом случае задача сравнительно легко может быть решена до конца. Здесь прежде всего интересно сопоставить количество энергии, падающей на дифрагирующую частицу, с энергией, ею рассеянной.

Рассмотрим этот вопрос в случае, когда рассеивающей частицей является, во-первых, линейный резонатор и, во-вторых, свободный электрон. Итак, рассмотрим линейный осциллятор (диполь, размеры которого гораздо меньше длины волны), на который падает волна от удаленного источника. Вычислим, как это обычно делают, по Герцу, количество энергии, рассеянное диполем [1].

---

[1] Заметим, что по существу нужно было бы более точно выяснить, что мы подразумеваем под «энергией», рассеянной диполем». Если взять любую поверхность, окружающую диполь, и учесть полное поле на ней (включая, конечно, и поле падающей волны), то, очевидно, поток через поверхность будет нуль. Если взять замкнутую поверхность, включающую источник света, то поток энергии через нее будет отличен от нуля, так как для поддержания колебаний источника затрачивается работа. Если внутри поверхности заключен и источник и рассеивающий диполь, то поток через поверхность будет равен сумме трех членов: первый дает излучение источника, второй дает как раз энергию, рассеянную диполем, третий же — интерференционный, он зависит от расстояния между источником и диполем и при усреднении по расстоянию даст нуль.

Уравнение движения осциллятора, колеблющегося под действием поля световой волны, такое:

$$m\ddot{p} - \frac{2e^2}{3c^3}\dddot{p} + kp = e^2E_0e^{int}. \tag{12}$$

Поле излученной осциллятором волны определяется величиной $\ddot{p}$, и энергия, излученная за единицу времени, равна

$$W = \frac{\ddot{p}_0^2}{3c^3}, \tag{13}$$

где $\ddot{p}_0$ — амплитуда $\ddot{p}$.

Нам нужно решение, дающее вынужденные колебания

$$p = p_0e^{i(nt+\varphi)}. \tag{14}$$

Будем предполагать, что имеет место *резонанс*, т. е. $n^2 = \frac{k}{m}$; тогда $\varphi = \frac{\pi}{2}$, и, как видно из уравнения (12),

$$\ddot{p}_0 = -\frac{3c^3}{2n}E_0. \tag{15}$$

Подставляя это значение в (13), найдем величину энергии, рассеянной за единицу времени настроенным в резонанс диполем

$$W = \frac{3c}{16\pi^2}\lambda^2E_0^2, \tag{16}$$

где λ — длина волны падающего света, $\lambda = \frac{2\pi c}{n}$. В падающей волне в отсутствие осциллятора за единицу времени через единицу площади проходит энергия, равная

$$\frac{c}{8\pi}E_0^2. \tag{17}$$

Сравнивая (16) и (17), мы можем сказать, что осциллятор выбирает энергию с площади сечения падающей волны, равной

$$\sigma = \frac{3}{2\pi}\lambda^2.$$

Таким образом, при резонансе величина этой площади рассеяния σ определяется только длиной волны и не имеет никакого отношения к размерам осциллятора (мы уже предположили, что его размеры очень малы по сравнению с длиной волны).

Это очень важный факт; учитывая его, можно, между прочим, понять взаимное влияние осцилляторов, находящихся на расстоянии порядка длины волны друг от друга. Не следует, однако, упускать из виду, что на основании сказанного раньше рассматривать поток энергии через площадку с размерами порядка длины волны, по существу, не имеет смысла, так что полученному выражению для σ нельзя приписывать особенно наглядного буквального содержания.

Перейдем теперь к рассеянию на свободном электроне. В этом случае в (12) нужно положить $k = 0$, член с $\dddot{p}$ для не слишком коротких волн очень мал по сравнению с членом с $\ddot{p}$ (если $\lambda \gg \frac{e^2}{mc^2}$), так что мы можем написать

$$\ddot{p}_0 = \frac{e^2}{m} E_0.$$

Рассеянная энергия согласно (13) равна

$$W = \frac{e^4}{3c^3m^2} E_0^2;$$

сравнение с (17) дает теперь для площади рассеяния

$$\sigma = \frac{8\pi e^4}{3c^4m^2}.$$

Полагая $\sigma = \pi R^2$, найдем для радиуса площади рассеяния величину

$$R = \sqrt{\frac{8}{3}} \frac{e^2}{mc^2},$$

отличающуюся только числовым фактором $\sqrt{\frac{8}{3}}$ от «радиуса электрона», равного $a = \frac{e^2}{mc^2}$. Таким образом, как заметил Борн, свободный электрон рассеивает столько, сколько получалось бы из простых геометрических расчетов, использующих «размер электрона». Борн идет обратным путем: он исходит из представления о фотонах, считает их радиус равным нулю, использует эти геометрические соображения и, сравнивая результат с приведенным выше расчетом по формуле Герца, получает радиус электрона. Конечно, большой убедительности в этом рассуждении Борна нет, хотя оно и интересно.

ПЯТАЯ ЛЕКЦИЯ

(7.I 1933 г.)

Решение дифракционных задач по методу Релея. Применение к некоторым частным вопросам

Вернемся к задачам теории дифракции и разберем метод, применявшийся Релеем для их решения. Этот метод вполне эквивалентен применению функции Грина, так как при его применении мы считаем поле (но не его производную) на экране заданным, и приводит к тождественным результатам. Поэтому методу Релея присущи и все недостатки метода функции Грина. Однако многие задачи он позволяет решать очень изящно.

Рассмотрим опять двухмерную дифракционную задачу и будем, как и прежде (стр. 16), считать, что экран с отверстиями (щелями) лежит в плоскости $z = 0$, образующие щелей параллельны оси y, поле E направлено по оси y, волна падает на экран нормально (в направлении оси z). Частным интегралом волнового уравнения

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial z^2} + k^2 E = 0$$

будет выражение

$$e^{i(ux+z\sqrt{k^2-u^2})},$$

где u *произвольно*. При $|u| < k$ это решение представляет собой плоскую волну с постоянной амплитудой, распространяющуюся под углом φ к оси z, причем

$$\operatorname{tg}\varphi = \frac{u}{\sqrt{k^2-u^2}}.$$

x
φ
z

Рис. 11

При $|u| > k$ это решение изображает волну с амплитудой, экспоненциально убывающей в направлении оси z, бегущую вдоль оси x и синусоидально зависящую от x

$$e^{iux-z\sqrt{u^2-k^2}}.$$

В силу линейности волнового уравнения его решением будет и сумма

$$E = \sum_m C_m e^{i(u_m x+z\sqrt{k^2-u_m^2})} \tag{18}$$

при любых C_m и u_m.

Рассмотрим случай периодической структуры экрана, когда щели в экране расположены по периодическому закону (дифракционная решетка). Строго говоря, этот случай выпадает из нашего прежнего рассмотрения, так как это бесконечная структура, однако этот вопрос мы пока оставляем в стороне [1]. Заметим, что в этом случае можно сформулировать общие положения, сохраняющиеся и в строгой теории и связанные только с периодичностью структуры экрана.

Можно ли подобрать C_m и u_m так, чтобы удовлетворить граничным условиям — заданному распределению поля на экране? Если это удастся, то в силу теоремы единственности для решений волнового уравнения полученное решение должно совпасть с решением по методу Кирхгофа с применением функции Грина. Мы увидим сейчас, что это действительно так.

Пусть при $z = 0$ задано распределение поля на решетке как периодическая функция x

$$E(x, 0) = f(x).$$

Тогда необходимо, следовательно, чтобы наше решение (18) обратилось при $z = 0$ в эту функцию $f(x)$

$$\sum_m C_m e^{iu_m x} = f(x).$$

Это осуществимо; достаточно положить

$$u_m = m\frac{2\pi}{d} = m\mu \qquad \left(\mu = \frac{2\pi}{d}\right),$$

где $m = 0, \pm 1, \pm 2, \ldots$, а d — период решетки, т. е. период функции $f(x)$. Мы получим тогда разложение функции в ряд Фурье, и коэффициенты этого ряда будут

$$C_m = \frac{1}{d}\int_{-\frac{d}{2}}^{+\frac{d}{2}} f(x)\, e^{-im\mu x}\, dx.$$

Таким образом, мы получили простой рецепт для решения задач о дифракции на бесконечных периодических структурах (решетках). Надо найти коэффициенты разложения Фурье C_m функции $f(x)$, представляющей поле на решетке, и решение запишется

---

[1] В частности, к этому случаю неприменимо обычное доказательство единственности решения волнового уравнения.

тогда в виде

$$E = \sum_{m=-\infty}^{+\infty} C_m e^{i(m\mu x + z\sqrt{k^2 - m^2\mu^2})}. \qquad (19)$$

Члены этого ряда для больших номеров m ($|m\mu| > k$) экспоненциально убывают в зависимости от z и на больших расстояниях от решетки поэтому не играют роли[1]. На больших расстояниях остаются только члены ряда, для которых $|m\mu| < k$. Это плоские волны постоянной амплитуды, распространяющиеся в направлениях, углы которых φ_m с осью z определяются условием

$$m\mu = k \sin\varphi_m \quad \text{или} \quad \sin\varphi_m = m\frac{\lambda}{d}. \qquad (19')$$

Это известная формула для угла дифракции на решетке. C_m дает поэтому амплитуду спектра m-го порядка.

«Синусоидальная решетка», для которой

$$f(x) = \cos \mu x$$

[или вообще $f(x) = \cos(\mu x + \gamma)$], дает только спектры первого порядка и не дает даже прямо прошедшей волны (спектра нулевого порядка). Такая решетка может быть осуществлена, только если свет при прохождении через решетку в определенных местах меняет свою фазу на π. Для решеток, состоящих из щелей и непрозрачных частей, прямо прошедшая волна всегда будет присутствовать, так как в соответствующем разложении Фурье будет постоянный член. Обычную штриховую решетку мы можем представлять себе как суперпозицию синусоидальных решеток с периодами, находящимися в целом отношении к основному, причем каждая из этих решеток дает лишь два дифракционных максимума, соответствующих $\pm m$-му члену ряда (19).

---

[1] Строго говоря, нужно показать, что убывает не только каждый такой член ряда, но и бесконечная их сумма. Это вытекает из следующей теоремы Абеля о рядах (см., например, Г у р с а. Курс анализа, § 166. Русск. изд. 1911, или В а л л е П у с с е н. Курс анализа, I. М.— Л., ГТТИ, 1933, § 387, стр. 429): если $S = \sum_{m=0}^{\infty} \varepsilon_m V_m$, суммы $S_m = \sum_{m=0}^{n} \varepsilon_m V_m$ ограниченны, $\lim_{m\to\infty} \varepsilon_m = 0$ и $\varepsilon_{m+1} < \varepsilon_m$, то $|S| < A\varepsilon_0$, где A — постоянная. В нашем случае нужно положить $V_m = C_m e^{im\mu x}$, $\varepsilon_m = e^{-z\sqrt{k^2 - m^2\mu^2}}$. Тогда получим, что $|S| < Ae^{-z\sqrt{m_1^2\mu^2 - k^2}}$, причем m_1 есть наименьшее m, для которого $|m\mu| > k$.

Если $\mu > k$, т. е. $d < \lambda$, то все члены ряда (19), кроме соответствующего $m = 0$, будут экспоненциально убывать в зависимости от z. На дальних расстояниях от решетки останется поэтому только прямо прошедший свет, изображаемый членом ряда, соответствующим значению $m = 0$. Это влечет за собой невозможность получить изображение решетки с периодом, меньшим длины волны, с помощью какого бы то ни было оптического прибора, например микроскопа. Мы получаем, таким образом, предел разрешения микроскопа, указанный Аббе и Гельмгольцем.

При образовании изображения решетки в оптическом приборе отпадают, с одной стороны, в силу только что сказанного все члены разложения поля на решетке, для которых $|m\mu| > k$: для них получается затухающее поле. С другой стороны, во всяком оптическом приборе, например в микроскопе, отрезаются благодаря конечной его апертуре высшие члены разложения Фурье поля решетки (вообще говоря, начиная с члена меньшего номера, чем благодаря первой причине). Если структура решетки грубая, такая, что уже первые члены разложения дают хорошее ее приближение, то аппарат хорошо ее изображает. Однако тонкие детали в изображении пропадают. Аппараты с «темным полем», например ультрамикроскоп, отрезают первые члены разложения, и в этом случае в изображении структура может совершенно исказиться, могут получиться мнимые детали.

Изложенные свойства дифракции на решетках можно применить к некоторым акустическим задачам. Заставим плоскую мембрану колебаться с определенной частотой. На ней образуются стоячие волны. Мы можем рассматривать такую мембрану как дифракционную решетку, считая, что акустическое поле в воздухе у поверхности мембраны задается ее колебаниями. Длина стоячей волны на мембране определяет период решетки. Если этот период очень мал — меньше длины акустической волны в воздухе, соответствующей данной частоте,— то распространения волн в воздухе не будет, а получится экспоненциально убывающее поле.

Этим объясняется опыт Лесли с колокольчиком (заменяющим мембрану), помещенным под колокол воздушного насоса. Если под колоколом воздух, то звон колокольчика слышен снаружи колокола. Если же воздух из-под колокола выкачать и напустить туда водород, то звон колокольчика слышен еще слабее, чем при разреженном воздухе. Здесь дело в том, что при той же частоте звука длина волны звука в водороде (из-за его малого молекулярного веса) гораздо больше, чем в воздухе, и больше, чем постоянная той «решетки», которая образуется на колокольчике при его колебаниях. В акустическом случае ясно видно, почему при этих условиях излучение падает. Если расстояние от узлов

до пучностей мембраны мало по сравнению с длиной волны звука в окружающей среде, то среда просто будет перетекать от пучностей к узлам (скорость этого течения при этих условиях меньше скорости звука). Это объяснение опыта Лесли принадлежит Стоксу [1]; до него этому опыту давали самые фантастические объяснения.

Применим теперь общие принципы релеевской теории дифракции к некоторым вопросам, намеченным еще вначале. Поставим следующий вопрос. Можно ли создать такую решетку, чтобы поле на ней равнялось

$$f(x) = e^{-iux}, \qquad u = \frac{2\pi}{d} \tag{20}$$

и, следовательно, чтобы эта решетка давала только *одну* дифрагированную волну? Такая решетка образуется при преломлении света на границе двух сред. Действительно, в первой среде поле равно сумме полей падающей и отраженной волн

$$E_1 = e^{-ik_1(x \sin\psi_1 + z\cos\psi_1)} + Ae^{-ik_1(x\sin\psi_1 + z\cos\psi_1)}$$

и, следовательно, на границе сред при $z = 0$

$$E_1 = (1 + A)\, e^{-ik_1 x \sin\psi_1}. \tag{21}$$

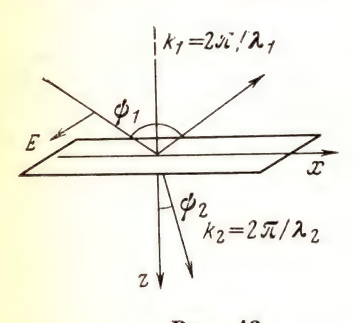


Рис. 12

Мы рассматриваем случай, когда электрический вектор перпендикулярен плоскости падения: $E_1 = E_y$ (см. рис. 12). В силу непрерывности тангенциальных компонент электрического вектора поле во второй среде у границы ($z = 0$) будет поэтому то же самое (21), но это как раз поле типа (20), причем $u = k_1 \sin\psi_1$. Поэтому во второй среде получается одна «дифрагированная волна» — это преломленная волна. Угол нормали ее с осью z определяется из уравнения

$$k_2 \sin\psi_2 = u = k_1 \sin\psi_1,$$

т. е. $\frac{\sin\psi_1}{\sin\psi_2} = \frac{\lambda_1}{\lambda_2}$. Мы получили закон преломления. Если $\sin\psi_1 > \lambda_1/\lambda_2$, то $u > k_2$ и во второй среде будет поле с амплитудой, убывающей в зависимости от z. Это случай полного внутреннего отражения. Поле во второй среде при этом получается того же

---

[1] *Релей*. Теория звука, т. II, § 324.

типа, что и от дифракционной решетки с очень мелкой структурой.

До сих пор мы делали выводы общего характера, исходя из того, что решение написано в виде суммы плоских волн. Нужно, однако, иметь в виду, что физический смысл имеет лишь их суперпозиция. Выделить отдельные плоские волны можно при помощи установленной на бесконечность трубы.

Рассматривая сумму плоских волн, дающую решение задачи для бесконечной решетки, мы ни на каких расстояниях от нее не получим картины с дифракционными максимумами в определенных местах.

Чтобы получить результаты, сравнимые с опытом, можно идти двумя путями. Во-первых, можно рассчитать, что получится, если за решеткой поставить определенным образом фокусированную трубу или просто линзу. Однако расчет этого случая не очень прост. Во-вторых, можно перейти от бесконечной решетки к рассмотрению ограниченных решеток. Применяя к этому случаю метод Релея, мы получим решение уже не в виде ряда Фурье (19), а в виде интеграла Фурье

$$E(x, z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(u)\, e^{i(ux + z\sqrt{\kappa^2 - u^2})} du. \tag{22}$$

При этом функцию $g(u)$ нужно определить из заданного распределения поля $E(x, 0) = f(x)$ в плоскости решетки (или экрана с отверстием) $z = 0$

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(u)\, e^{iux} du$$

так, что

$$g(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)\, e^{-iux} dx.$$

В этом случае для очень больших z можно приближенно найти значение интеграла (22) и получить распределение поля на очень больших расстояниях от решетки. Распределение поля там будет такое же, как и в фокальной плоскости линзы, поставленной за решеткой.

При дифракции на щели ширины $2a$ в экране и при нормальном падении волны

$$f(x) = 1, \qquad |x| < a,$$
$$f(x) = 0, \qquad |x| > a.$$

Поэтому

$$g(u) = \frac{1}{\sqrt{2\pi}} \int_{-a}^{+a} e^{-iux} dx = -\frac{2a}{\sqrt{\pi}} \frac{\sin ua}{ua}.$$

При дифракции на решетке периода $d = 2D$, состоящей из M щелей,

$$g(u) = \frac{1}{\sqrt{2\pi}} \sum_{m=-M/2}^{+M/2} e^{-imu} \int_{-a}^{+a} e^{-iux} dx = \frac{1}{\sqrt{2\pi}} \frac{\sin au}{au} \frac{\sin MDu}{\sin Du},$$

где a — ширина каждой щели.

ДОБАВЛЕНИЕ К ПЯТОЙ ЛЕКЦИИ

[*Строгое решение задачи о дифракции на очень узкой щели*]

Как мы знаем (лекции 3 и 4), если щель невелика по сравнению с длиной волны, то решение дифракционной задачи по методу Кирхгофа дает неправильный результат. В предельном случае, когда ширина щели очень мала по сравнению с длиной волны, удается, однако, дать *строгое* решение задачи.

Мы рассматривали плоский, бесконечно тонкий (находящийся в плоскости $z = 0$) экран из идеально проводящего материала. Щель в нем бесконечной длины, параллельна оси y и ширина ее $2a$. Со стороны отрицательных z на экран нормально падает волна, электрический вектор которой $E = E_y$ параллелен щели.

В этом случае задача строгой теории состоит в том, чтобы найти решение $E\ (x, z)$ волнового уравнения

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial z^2} + k^2 E = 0,$$

удовлетворяющее следующим условиям:

1) E и $\frac{\partial E}{\partial z} = ikH$ всюду вне экрана, включая и место щели, непрерывны;

2) на экране при $z = 0$ и $|x| > a$ $E = 0$ (граничное условие на идеальном проводнике);

3) при $z \to -\infty$ (со стороны падающей волны) поле складывается из падающей волны e^{ikz} и волн, идущих от экрана. Для построения этого решения положим

$$E = e^{ikz} - e^{-ikz} + \varphi(x, z).$$

Тогда, как легко убедиться, из условий для E вытекает, что $\varphi(x, z)$ будет решением волнового уравнения, удовлетворяющим следующим условиям:

1) всюду вне плоскости $z = 0$, φ и $\frac{\partial\varphi}{\partial z}$ непрерывны, но на щели при $z = 0$, $|x| < a$, φ остается непрерывным, а $\frac{\partial\varphi}{\partial z}$ испытывает скачок: $\left[\frac{\partial\varphi}{\partial z}\right]_{z=-0}^{z=+0} = 2ik$;

2) всюду на экране, т. е. при $z = 0$, $|x| > a$, φ равно нулю, так что если положить $\varphi(x, 0) = f(x)$, то $f(x) = 0$ при $|x| > a$;

3) φ стремится к нулю при $z \to \pm\infty$ и изображает волны, идущие в обе стороны от щели.

Решение задачи будем искать в виде, аналогичном (22), т. е.

$$\varphi(x, z) = \int_{-\infty}^{+\infty} g(u)\, e^{i(ux \pm z\sqrt{k^2 - u^2})} du. \tag{23}$$

Это решение волнового уравнения. Взяв перед корнем знак плюс для $z > 0$ и знак минус для $z < 0$, мы удовлетворим условию 3).

Полагая $z = 0$, находим

$$\varphi(x, 0) = f(x) = \int_{-\infty}^{+\infty} g(u)\, e^{iux} du,$$

так что

$$g(u) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(x)\, e^{-iux} dx.$$

Принимая во внимание условие 2), получим

$$g(u) = \frac{1}{2\pi} \int_{-a}^{+a} f(x)\, e^{-iux} dx. \tag{24}$$

Остается удовлетворить условию 1). Для этого нужно взять подходящую функцию $f(x)$ $(|x| < a)$, и в этом заключается вся трудность задачи.

Заметим, что φ, выражающееся равенством (23), очевидно, непрерывно при $z=0$ при любой $g(u)$, любой $f(x)$. Нужно, следовательно, удовлетворить еще только условию для скачка производной $\frac{\partial\varphi}{\partial z}$ при $z=0$. Дифференцируя (23), получим

$$\left[\frac{\partial\varphi}{\partial z}\right]_{z=-0}^{z=+0} = 2i\int\limits_{-\infty}^{+\infty} g(u)\sqrt{k^2-u^2}e^{iux}du.$$

Следовательно, условие скачка для $\frac{\partial\varphi}{\partial z}$ дает уравнение

$$\frac{1}{k}\int\limits_{-\infty}^{+\infty} g(u)\sqrt{k^2-u^2}e^{iux}du = 1, \qquad |x|<a, \tag{25}$$

причем $g(u)$ выражается через $f(x)$ согласно (24).

До сих пор рассуждения были общими, предположения о том, что щель узкая, сделано не было. Покажем теперь, что для *узкой щели*, для которой $a\ll\lambda$, $ka\ll 1$, мы удовлетворим уравнению (25), если возьмем [1]

$$f(x)=\frac{k}{i}\sqrt{a^2-x^2}.$$

При такой функции $f(x)$, $g(u)$ равна

$$g(u)=\frac{k}{2\pi i}\int\limits_{-a}^{+a}\sqrt{a^2-x^2}e^{-iux}dx=\frac{ka^2}{2\pi i}\int\limits_{-1}^{+1}\sqrt{1-\mu^2}e^{-iua\mu}d\mu=$$

$$=\frac{ka^2}{\pi i}\int\limits_{0}^{1}\sqrt{1-\mu^2}\cos(ua\mu)\,d\mu.$$

Учитывая, что [2]

$$\int\limits_0^1\sqrt{1-\mu^2}\cos\alpha\mu\,d\mu=\frac{\pi}{2}\,\frac{J_1(\alpha)}{\alpha},$$

где $J_1(\alpha)$ — бесселева функция первого порядка, получим

$$g(u)=\frac{ka^2}{2i}\,\frac{J_1(au)}{au}. \tag{26}$$

---

[1] Это распределение поля на узкой щели получено в работе Релея (Sci. Pap., т. IV, стр. 282 и 305) из других соображений.

[2] См., например, *Р. Курант, Д. Гильберт.* Методы математической физики. М.— Л., ГТТИ, 1933, т. I, гл. VII, § 2, формула (19) или *Кузьмин.* Бесселевы функции. ОНТИ, 1935, § 6, стр. 89.

Покажем теперь, что это выражение для $g(u)$ удовлетворяет уравнению (25) с точностью до членов порядка ka. Подставляя (26) в (25), получаем

$$\frac{1}{k}\int_{-\infty}^{+\infty} g(u)\sqrt{k^2-u^2}e^{iux}du = \frac{a}{2i}\int_{-\infty}^{+\infty}\frac{J_1(au)}{u}\sqrt{k^2-u^2}e^{iux}du =$$

$$= \frac{a}{i}\int_0^{\infty}\frac{J_1(au)}{u}\sqrt{k^2-u^2}\cos ux\,du = I_1+I_2+I_3,$$

где

$$I_1 = a\int_0^{\infty} J_1(au)\cos ux\,du,$$

$$I_2 = a\int_0^{k} J_1(au)\cos ux\cdot\left[\frac{\sqrt{k^2-u^2}}{iu}-1\right]du,$$

$$I_3 = a\int_0^{\infty} J_1(au)\cos ux\cdot\left[\frac{\sqrt{u^2-k^2}}{u}-1\right]du.$$

Первый интеграл можно взять точно, если воспользоваться известными формулами [1] для интегралов, содержащих функции Бесселя,

$$\int_0^{\infty} J_1(\xi)\cos\xi y\,d\xi = \begin{cases} 1 & \text{при } |y|<1,\\ 1-\dfrac{1}{\sqrt{y^2-1}} & \text{при } |y|>1.\end{cases}$$

Таким образом,

$$I_1 = 1.$$

Второй и третий интегралы мы оценим приближенно. Так как для $|u|<k$

$$\left|\frac{\sqrt{k^2-u^2}}{iu}-1\right| = \left|\frac{\sqrt{k^2-u^2}-iu}{u}\right| = \frac{k}{u} \quad \text{и} \quad |J_1(au)| < \text{const}\cdot au,$$

то

$$|I_2| < a\int_0^{k}|J_1(au)|\,|\cos ux|\left|\frac{\sqrt{k^2-u^2}}{iu}-1\right|du <$$

$$< ka^2\cdot\text{const}\int_0^{k}du = \text{const}\cdot k^2u^2.$$

[1] См. *Кузьмин*. Бесселевы функции, гл. IV, § 4.

Для оценки I_3 учтем, что при $|u| > k$

$$\left|\frac{\sqrt{u^2-k^2}}{u} - 1\right| = \frac{k^2}{u\,(\sqrt{u^2-k^2}+u)} \leqslant \frac{k^2}{u^2} \quad \text{и} \quad |J_1(au)| < 1,$$

так что

$$|I_3| < k^2 a \int\limits_k^{\infty} \frac{du}{u^2} = ka.$$

Таким образом,

$$\frac{a}{i}\int\limits_{-\infty}^{+\infty} \frac{J_1(au)}{u} \sqrt{k^2-u^2}\, e^{iux} du = 1 + O(ka),$$

и выражение (26) удовлетворяет уравнению (25) с точностью до членов порядка не ниже ka. Мы видим, таким образом, что в то время как для очень широкой щели

$$g(u) = \frac{1}{\pi}\,\frac{\sin au}{u}, \tag{27}$$

для щели очень узкой по сравнению с длиной волны это выражение нужно заменить на

$$g(u) = \frac{ka}{2i}\,\frac{J_1(ku)}{u}. \tag{26'}$$

Это выражение отличается в максимуме в $\frac{\pi}{4}ka$ раз от значения, которое имеет в максимуме выражение (27) для широкой щели. Фазы их отличаются на $\frac{\pi}{2}$, и, кроме того, (26) несколько быстрее убывает при возрастании u, чем (27).

ШЕСТАЯ ЛЕКЦИЯ

(13. II 1933 г.)

Соотношение неопределенностей в квантовой механике. Прохождение электронов через щель. Соотношение между степенью монохроматичности колебания и его продолжительностью и его роль в радиотелефонии

Разберем теперь парадокс, связанный с «соотношением неопределенностей» квантовой механики.

Пусть мы имеем волновую функцию частицы

$$\psi = Ae^{i(ux-nt)}, \tag{28}$$

удовлетворяющую уравнению Шрёдингера при свободном движении частицы. При этом n и u связаны с энергией E и импульсом частицы p соотношениями

$$n = \frac{E}{\hbar}, \qquad u = \frac{p}{\hbar}, \qquad \hbar = \frac{h}{2\pi}. \tag{28'}$$

Величина $|\psi|^2\, dx$ дает относительную вероятность найти частицу в интервале x, $x + dx$. В нашем случае эта вероятность не зависит от x, координата частицы совершенно неопределенна. Наоборот, импульс имеет вполне определенное знечение $p = \hbar u$.

Это частный случай соотношения неопределенностей. Вообще говоря, вероятность положения частицы дается выражением $|\psi|^2\, dx$, вероятность же того, что частица имеет импульс, лежащий между p и $p + dp$, дается выражением

$$\frac{1}{\hbar}\left| g\left(\frac{p}{\hbar}\right)\right|^2 dp, \tag{29}$$

где $g\,(u)$ — функция, входящая в разложение Фурье функции $\psi\,(x)$

$$\psi\,(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} g\,(u)\, e^{iux} du. \tag{30}$$

«Принцип неопределенности» вытекает тогда из того, что чем «острее» функция $\psi\,(x)$, тем «расплывчатее» $g\,(u)$, и наоборот. Количественно этот принцип выражается «соотношением неопределенностей»

$$\Delta p \cdot \Delta x \geqslant \frac{\hbar}{2}, \tag{31}$$

где Δp — неопределенность в импульсе, а Δx — неопределенность в положении. Существование соотношения (31) можно показать

на ряде примеров. Один из этих примеров, который часто приводят, следующий. Поставим на пути электронов диафрагму, т. е. экран с отверстием (рис. 13).

В отсутствие диафрагмы компонента импульса по оси x вполне определенна: $p_x = 0$, но координата x совершенно неопределенна. При наличии диафрагмы координата x определяется с точностью до ширины отверстия $\Delta x = 2a$, но теперь будет неопределенность и в p_x, так как благодаря дифракции электроны отклоняются в стороны. При этом говорят, что будет выполняться соотношение

$$\Delta p_x \cdot \Delta x \geqslant \frac{\hbar}{2}.$$

Но мы уже видели (лекция 1), что это как будто не так. Действительно, максимальная неопределенность в компоненте импульса Δp_x которая может получиться благодаря вызванному дифракцией отклонению электронов, будет меньше p_x и, значит, меньше величины полного импульса p_0

$$\Delta p_x \leqslant p_0.$$

Следовательно,

$$\Delta p_x \cdot \Delta x \leqslant 2ap_0,$$

а эту величину можно сделать сколь угодно малой, уменьшая ширину отверстия $2a$.

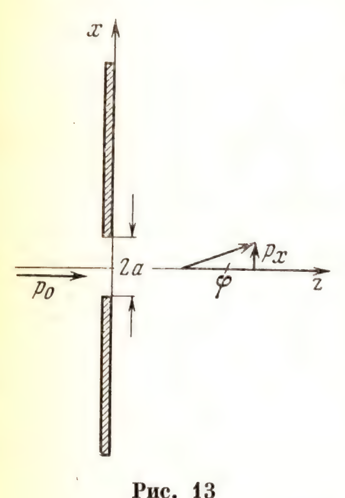


Рис. 13

В чем здесь дело? Ответ в том, что при малом a за отверстием будут не только распространяющиеся, но и экспоненциально убывающие волны. Если при рассмотрении дифракции принять их во внимание, то противоречие исчезает.

Разберем вопросы, связанные с «соотношением неопределенностей» в волновой механике, более подробно. Уравнение Шрёдингера для одной степени свободы имеет вид

$$\frac{\partial^2\psi}{\partial x^2} - \frac{2m}{\hbar^2} V(x)\psi = \frac{2im}{\hbar} \frac{\partial\psi}{\partial t}. \tag{32}$$

Как и во всякой физической теории, в квантовой механике прежде всего возникает вопрос о том, как переходить от математических символов к физическим величинам. Мы должны указать, что означают в уравнении (32) входящие в него буквы x, t,

ψ и т. д. Физический смысл ψ был выяснен не сразу, но окончательное ее толкование — статистическое.

Каждой опытной установке соответствует определенное уравнение Шрёдингера (например, для свободно движущейся точки $V = 0$) и определенное его решение. Волновая механика не дает, вообще говоря, ответов, относящихся к единичному опыту. Функция $\psi(t, x)$ (известная для какого-то момента t) относится не к данному опыту, а к данному опытному устройству, к серии опытов с данным устройством, которое характеризуется видом $V(x)$.

Выражение $\psi\psi^* dx$ дает вероятность найти частицу в момент t в точке с координатой между x и $x + dx$. При этом, чтобы иметь дело с обычными вероятностями, нужно нормировать так, что $\int\limits_{-\infty}^{+\infty} \psi\psi^* dx = 1$. Последнее возможно, если этот интеграл сходится (это бывает не всегда; в частности, в случае свободной частицы это не так).

Но задания положения электрона еще недостаточно для характеристики его состояния, надо знать еще его импульс. Величину импульса волновая механика характеризует тоже только статистически. Для этого она дает следующий рецепт [который мы уже указывали — формула (30)]. Нужно разложить $\psi(x)$ (при заданном t) в интеграл Фурье

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} g(u)\, e^{iux}\, dx.$$

Тогда

$$\frac{1}{\hbar} \left| g\left(\frac{p}{\hbar}\right) \right|^2 dp$$

дает вероятность того, что импульс (в этот момент t) лежит между p и $p + dp$.

Заметим, что с точки зрения теории относительности этот рецепт очень естествен. Для свободной частицы $V = 0$, и решение уравнения Шрёдингера имеет вид (28), где n играет роль циклической частоты, а u — волнового числа. Соотношение $E = \hbar n$ было еще у Планка. Де Бройль перенял его и ввел в волновую механику. С точки зрения теории относительности энергия и три компоненты импульса (умноженные на c) образуют четырехмерный вектор, совершенно так же, как и частота и три компоненты волнового вектора. Поскольку между t-компонентами обоих четырехмерных векторов принята связь $E = \hbar n$, то такая же связь должна иметь место и для компонент по пространственным осям.

Это и записано для одного пространственного измерения в соотношении $p = \hbar u$ (28′). Если $\psi(t, x)$ не является синусоидальной волной вида (28), то, применяя теорему Фурье, мы представляем ψ как суперпозицию таких волн и амплитуду, с которой представлена каждая такая волна, т. е. $g(u)$ естественно считать мерой вероятности того, что импульс имеет значение $p = \hbar u$.

Эти основные положения о вероятности координаты и вероятности импульса влекут за собой чрезвычайно важные следствия. Именно, исходя из свойств интеграла Фурье, можно показать, что если кривая, изображающая $|\psi|^2$ в зависимости от x, имеет очень острый максимум, то кривая, изображающая $|g|^2$ в зависимости от u, очень широкая, расплывчатая. Это значит, что очень точное определение координаты несовместимо с точным определением импульса, и наоборот. Невозможен такой опыт, с помощью которого измеряются абсолютно точно и координата и импульс.

Чтобы это качественное утверждение формулировать в количественной форме, нужно ввести меры разброса значений координат и импульса. Для этого поступают так.

Определяют средние значения величин x и u

$$\bar{x} = \int_{-\infty}^{+\infty} x\,|\psi|^2\,dx, \qquad \bar{u} = \int_{-\infty}^{+\infty} u\,|g(u)|^2\,du$$

и величины Δx и Δu, определяемые равенствами

$$(\Delta x)^2 = \int_{-\infty}^{+\infty} (x-\bar{x})^2\,|\psi|^2\,dx, \qquad (\Delta u)^2 = \int_{-\infty}^{+\infty} (u-\bar{u})^2\,|g|^2\,du, \qquad (33)$$

рассматривают как меры неточности (меры разброса) величин x и u. Конечно, в этом определении меры разбросов есть известный произвол.

Для определенных указанным образом величин Δx и Δu имеет место чисто математическая теорема, заключающаяся в следующем. Если ψ и g — любые функции, связанные между собой по Фурье [согласно (30)], то Δx и Δu, связанные с ψ и g согласно (33), удовлетворяют неравенству

$$\Delta x \cdot \Delta u \geqslant \frac{1}{2}.$$

Для Δx и $\Delta p = \hbar \Delta u$ отсюда вытекает соотношение

$$\Delta x \cdot \Delta p \geqslant \frac{\hbar}{2}.$$

Это и есть количественная формулировка «соотношения неопределенностей». Знак равенства имеет в нем место только в том случае, если $|\psi|^2$ изображается гауссовой кривой. Из этого соотношения сразу же видно, что если, например, $\Delta x = 0$, т. е. частица фиксирована в определенном месте, то $\Delta p = \infty$. Это значит, что при этом мы так воздействуем на частицу, что разброс величины ее импульса оказывается бесконечно большим.

Обычно, рассматривая примеры соотношения неопределенностей, оперируя с «разбросами» координаты δx и импульса δp, не дают точного их определения, а оценивают их «на глаз» и показывают, что для них

$$\delta x \cdot \delta p > \hbar.$$

Для двух степеней свободы дело обстоит совершенно так же. Здесь волновая функция есть $\psi(x, z)$. Вероятность найти координаты частицы лежащими в пределах x, $x + dx$ и z, $z + dz$ равна

$$|\psi|^2 \, dx \, dz.$$

Для определения вероятности компонент импульса надо разложить $\psi(x, z)$ в интеграл Фурье

$$\varphi(x, z) = \frac{1}{2\pi} \iint\limits_{-\infty}^{+\infty} g(u, v) \, e^{i(ux+vz)} \, du \, dv,$$

и тогда

$$|g(u, v)|^2 \, du \, dv = \frac{1}{\hbar^2} \left| g\left(\frac{px}{\hbar}, \frac{pz}{\hbar}\right) \right|^2 dp_x dp_z$$

дает вероятность того, что компоненты импульса лежат в пределах p_x, $p_x + dp_x$; p_z, $p_z + dp_z$. Аналогично обстоит дело при большем числе измерений.

Обратимся теперь к опыту с дифракцией электронов, проходящих через диафрагму, о котором мы уже говорили раньше. На щель в экране падает пучок электронов с вполне определенным импульсом (см. рис. 13)

$$p_x = p_y = 0, \qquad p_z = k\hbar = \frac{2\pi\hbar}{\lambda} = p_0.$$

Волновая функция, соответствующая этим электронам, имеет вид

$$\psi = e^{ikz}.$$

До диафрагмы координата x неопределенна, разброс ее бесконечно большой, как это и дожно быть, поскольку p_x точно за-

дана. После прохождения электронов через диафрагму (имеющую вид щели ширины $2a$, параллельной оси x) координата x определена с точностью $\delta x \sim a$, разброс p_x не может быть больше p_z, и, таким образом, $\delta x \cdot \delta p_x \leqslant a p_z$ в противоречие с соотношением неопределенностей.

Рассмотрим, однако, вопрос более точно. Какие будут волны за щелью?

Ответ нам дает решение задачи о дифракции на щели. Применяя метод Релея (лекция 5), получим решение, дающее распределение ψ за экраном, в виде

$$\psi(x, z) = \text{const} \int_{-\infty}^{+\infty} \frac{\sin au}{au} e^{i(ux + z\sqrt{k^2 - u^2})} du. \tag{34}$$

Здесь ψ представлена в виде суперпозиции волн вида $e^{i(ux+z\sqrt{k^2-u^2})}$. Для той части интеграла, для которой $|u| < k$, это — плоские волны постоянной амплитуды, распространяющиеся под углом φ к оси z, причем

$$u = k \sin \varphi = \frac{p_x}{\hbar}, \qquad \sqrt{k^2 - u^2} = k \cos \varphi = \frac{p_z}{\hbar}.$$

Амплитуда волн зависит от u по закону $\frac{\sin au}{au}$, и, следовательно, вероятность значения импульса, близкого к $p_x = \hbar u$, пропорциональная квадрату этого выражения, будет иметь заметную величину только при условии, что по порядку величины $au \sim 1$, или $ka \sin \varphi \sim 1$, т. е.

$$\frac{p_x a}{\hbar} \sim 1.$$

Значения p_x такого порядка и меньшие по абсолютной величине будут иметь еще заметную вероятность. Разброс импульса δp_x будет равен этому значению p_x ($\delta p_x \sim p_x$), так как среднее значение $\overline{p}_x = 0$. Разброс x-координаты можно считать равным a, $\delta x \sim a$.

Таким образом,

$$\delta x \cdot \delta p_x \sim \hbar.$$

Это мы и хотели показать. Однако при выводе мы ограничились волнами, для которых $|u| < k$, т. е. волнами, для которых $p_x = \hbar u < \hbar k = p_0$. Поэтому наш вывод верен, только если $\delta p_x < p_0$. Это ограничение для разброса δp_x и было принято при качественном выводе, указанном в начале лекции и приводящем, как мы видели, при достаточно узкой щели к противоречию с соотношением неопределенностей.

Противоречие разрешается, если принять во внимание, что мы не учли волн, для которых $|u| > k$.

Если ширина щели a мала (невелика по сравнению с дебройлевской длиной волны $\lambda = \hbar/k$), то величина $\left|\frac{\sin au}{au}\right|$ еще велика при больших u (для которых $|u| > k$) и членами интеграла с этими u пренебрегать нельзя. В этом случае вероятность того, что компонента импульса p_x больше $\hbar/a$, значительна и $\delta p_x > \hbar/a$.

Для этих u ($|u| > k$) элементы интеграла представляют собой «затухающие» волны, т. е. волны с амплитудой, убывающей в направлении оси z. Для таких волн уже нельзя говорить об определенном полном импульсе (только плоская волна постоянной амплитуды изображает состояние с точно определенным полным импульсом). Поэтому вероятность того, что p_x больше полного импульса p_0, отлична от нуля и может быть значительна; неравенство $\delta p_x < p_0$ не имеет места.

Неравенство $\delta p_x \cdot \delta x > \hbar$ сохраняется, таким образом, и при малой ширине щели.

Парадокс получился потому, что за экраном предполагалось наличие только синусоидальных плоских волн *постоянной* амплитуды, имеющих разные направления, но одинаковое волновое число k, в то время как при малом a там существенно представлены и затухающие плоские волны. Это значит, что за экраном полный импульс может быть больше, чем его значение p_0 до экрана.

Представление волновой механики об импульсе не укладывается в представления обычной классической механики о скорости и импульсе. В качестве примера возьмем электрон, движущийся в ящике (в одном измерении), отражаясь от его стенок. По классической механике его скорость (а значит, и импульс) может иметь только два значения: $\pm v$. По волновой механике у электрона при данной энергии бесконечно много возможных импульсов (и скоростей), вплоть до бесконечно больших значений импульса.

Еще одно замечание. Наши выводы сделаны на основании дифракционной формулы (34), выведенной в предположении, что поле на экране задано. В волновой механике, как и в волновой оптике, при узкой щели задавать поле на экране нельзя. К счастью, для очень узких щелей нужную здесь задачу волновой механики можно решить строго (см. добавление к лекции 5). При этом оказывается, что получается формула, отличающаяся от (34) только заменой $\frac{\sin au}{au}$ на $\frac{J_1(au)}{au}$, где J_1 — бесселева функция первого порядка. Это значит, что все сделанные нами выводы сохраняются, так как в нужном для нас отношении эти две функции ведут себя аналогично.

Перейдем теперь к более прозаическим вопросам. Графическая сторона соотношения неопределенностей состоит, как уже указывалось, в том, что чем у́же кривая, тем больше нужно синусоид, чтоб ее представить. Чем резче локализована волна в пространстве, тем менее она монохроматична, тем больший участок захватывает ее спектр. Это относится к распределению волны в пространстве. Аналогичное положение справедливо и для зависимости от времени. Короткий импульс — острая локализация во времени — несовместим с узким спектром частот, и обратно. Это обстоятельство имеет первостепенное практическое значение в радиотелеграфии. Путем настройки, применяя острый резонанс, мы защищаемся от посторонних станций и помех. Чем острее настройка приемника, тем меньше будут мешать посторонние станции, Но такой остроселективный приемник не может принимать коротких сигналов, так как короткий сигнал имеет широкий спектр. Между тем применение коротких сигналов нужно для того, чтобы можно было быстро телеграфировать. Значит, при быстром телеграфировании приходится брать приемник с широкой полосой пропускания, а тогда ему будут мешать посторонние станции. Здесь, по существу, играет роль то же соотношение, что и в принципе неопределенности. Поэтому людям, знающим радиотелеграфию, принцип неопределенности бывает легко объяснить.

В формулировке этого принципа для бегущих волн в пространстве, с одной стороны, и для колебаний во времени — с другой, есть интересное различие. Мы называем пространственную волну (в одном измерении) монохроматической, если ее зависимость от координаты дается функцией Ce^{iux}, где C — комплексная постоянная. Волны, изображаемые функциями e^{iux} и e^{-iux}, — разные волны; это волны, идущие в противоположных направлениях. В этом смысле

$$\cos ux = \frac{1}{2} e^{iux} + \frac{1}{2} e^{-iux}$$

— волна не монохроматическая, а суперпозиция двух разных монохроматических волн.

При рассмотрении колебаний во времени монохроматическим мы называем колебание вида $A \cos \omega t + B \sin \omega t$, где A и B — произвольные постоянные. В частности, колебание, изображаемое функцией $\cos \omega t$,— монохроматическое колебание. Это различие в том, что́ мы понимаем в этих случаях под «монохроматичностью», ведет к некоторым различиям в точной математической формулировке рассматриваемого общего положения [1].

---

[1] См. А. *Майер*, *Е. Леонтович*. ДАН СССР, 4, 353, 1934.

СЕДЬМАЯ ЛЕКЦИЯ

(25.III 1933 г.)

Явление Допплера при движении дифракционной решетки. Поле при дифракции на ограниченной решетке на больших расстояниях от нее. Аберрация при наличии дисперсии в эфире

Разберем теперь один парадокс, связанный с принципом Допплера. Рассмотрим дифракцию световой волны на дифракционной решетке. Предположим, что источник света, падающего на решетку, движется по направлению к ней.

Будем сначала рассуждать с точки зрения представления о неподвижном эфире и допустим, что решетка и наблюдатель покоятся по отношению к эфиру, а источник движется относительно него. Тогда частота падающего на решетку света равна

$$\nu = \nu_0 (1 + \beta),$$

а длина волны получится соответственно уменьшенной

$$\lambda = \frac{c_0}{\nu} = \frac{\lambda_0}{1 + \beta}, \tag{35}$$

где λ_0 — длина волны покоящегося источника, c_0 — скорость света относительно эфира, а β — отношение скорости источника к скорости света c_0.

Пользуясь дифракционной решеткой, мы измеряем длину волны.

Измеряя угол отклонения φ для первого дифракционного максимума, мы определим длину волны падающего света λ, пользуясь формулой $\lambda = d \sin \varphi$, и сможем заметить допплеровское смещение по отношению к опыту с неподвижным источником. Обратный случай, когда движется наблюдатель с решеткой, а источник покоится, с точки зрения принципа относительности, конечно, совершенно эквивалентен рассмотренному. С точки зрения же старой лоренцовой теории неподвижного эфира эти случаи различны. Как показал Лоренц, удерживая только члены первого порядка относительно β, электронная теория приводит к тем же выводам, что и теория относительности [1]. Между тем непосредственный расчет второго случая, который я сейчас приведу, как будто бы говорит о том, что это не так, и мы получаем здесь, таким образом, противоречие — парадокс.

---

[1] См., например, *Г. А. Лоренц.* Теория электронов, гл. V (примечание 72).

Действительно, посмотрим, как с точки зрения неподвижного эфира обстоит дело во втором случае. Частота света, падающего на движущуюся решетку от неподвижного источника, с точностью до членов первого порядка по-прежнем удается выражением $\nu = \nu_0 (1 + \beta)$. Но для дифракции существенна длина волны, а чтобы ее найти, нужно знать скорость света относительно движущейся решетки. Согласно представлениям теории неподвижного эфира, эта скорость получится применением теоремы о сложении скоростей и, очевидно, с точностью до первого порядка относительно β будет равна

$$c = c_0 (1 + \beta \cos\psi). \tag{36}$$

Здесь ψ — угол между направлением движения и направлением распространения света (рис. 14). Поскольку мы считаем, что решетка движется к источнику и плоскость решетки перпендикулярна к направлению движения, постольку для дифрагированного луча ψ, очевидно, и есть угол дифракции. Длина волны дифрагированного света равна теперь, согласно (35) и (36),

$$\lambda = \frac{c}{\nu} = \lambda_0 \frac{1 + \beta\cos\psi}{1 + \beta} \approx \approx \lambda_0 [1 + \beta(\cos\psi - 1)]. \tag{37}$$

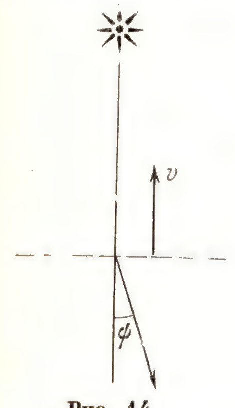


Рис. 14

Это выражение существенно отличается от (35). Из него вытекает, в частности, что в противоположность первому случаю для малых углов дифракции допплеровское смещение исчезает, $\lambda \approx \lambda_0$, и движения в этом случае нельзя будет обнаружить.

В чем ошибка? Ошибка последнего рассуждения заключается в следующем. В нем предполагается, что длина волны одинакова для всех лучей, в результате интерференции которых получается дифракционный максимум. Чтобы выяснить вопрос, рассмотрим подробнее образование дифракционной картины.

Практически дифракционную картину получают в фокальной плоскости линзы. Для теоретического разбора, однако, удобнее рассмотреть картину, получающуюся без линзы, но на очень далеком расстоянии от решетки, что мы и сделаем. Расчет дифракции с линзой в нашем случае очень сложен: нужно было бы учесть все изменения хода лучей, вносимые линзой, и различие длин волн в лучах разного наклона. Этого не стоит делать, так как

мы знаем, что картина на очень далекой плоскости должна быть такая же, как в фокальной плоскости линзы.

Предположим сначала, что решетка имеет только две щели (рис. 15). Расстояния точки наблюдения от щелей равны r и r'. При увеличении расстояния разность углов ψ и ψ', а значит, и разность длин волн λ и λ', идущих от первой и второй щели, будет уменьшаться. Однако число волн, укладывающихся на расстояниях r и r', будет становиться все больше.

Поэтому маленькая разница в длинах волн в результате приводит к такой разности хода между волнами от первой и второй щели, с которой необходимо считаться. Действительно, эта разность хода равна

$$\Delta = \frac{r}{\lambda} - \frac{r'}{\lambda'} = \frac{1+\beta}{\lambda_0}\left[\frac{r}{1+\beta\cos\psi} - \frac{r'}{1+\beta\cos\psi'}\right] \approx$$

$$\approx \frac{1+\beta}{\lambda_0}\left[r - r' - \beta\left(r\cos\psi - r'\cos\psi'\right)\right].$$

В дифракционном максимуме разность Δ должна равняться целому числу n. В первом приближении (пренебрегая, как всюду, квадратами β) это условие дает

$$r - r' - \beta\left(r\cos\psi - r'\cos\psi'\right) = \frac{n\lambda_0}{1+\beta}.$$

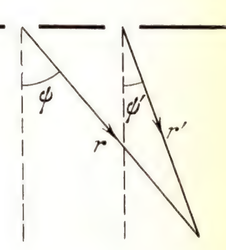


Рис. 15

При прежнем выводе мы считали, что $\psi = \psi'$. Это и привело нас к парадоксу, так как тогда (учитывая, что $r - r'$ нужно положить равным $d\sin\psi$), мы получаем прежний результат (37)

$$r - r' = d\sin\psi = \frac{n\lambda_0}{1+\beta}\,\frac{1}{1-\beta\cos\psi} \approx$$

$$\approx n\lambda_0\left[1 + \beta\left(\cos\psi - 1\right)\right].$$

В действительности же $\psi \neq \psi'$, а, как ясно из чертежа, $r\cos\psi = r'\cos\psi'$, так что

$$r - r' = d\sin\psi = \frac{n\lambda_0}{1+\beta}.$$

Мы измерим, следовательно, ту же длину волны, что и в случае движения источника.

Рассмотрим теперь *периодическую* и, следовательно, бесконечную решетку. Если на такую решетку нормально падает волна, то поле за решеткой, пользуясь методом Релея (лекция 5),

можно представить так:

$$E = \sum_n C_n e^{-i\left(\frac{2\pi n}{d}x + \sqrt{k^2 - \left(\frac{2\pi n}{d}\right)^2} z\right)}, \tag{38}$$

где $k = \frac{2\pi}{\lambda}$, а d — период решетки. Это суперпозиция плоских волн, распространяющихся в направлениях, определенных углами $\psi_n \left(\sin\psi_n = \frac{n\lambda}{d}\right)$, и затухающих волн. Если решетка движется, то длина волны λ выражается, согласно (37), как $\lambda \approx \lambda_0 [1 + \beta(\cos\psi - 1)]$, и, следовательно, для тех из совокупности входящих в (38) волн, которые распространяются под малыми углами, мы имеем опять (пренебрегая ψ^2) $\lambda = \lambda_0$. Таким образом, получается прежний парадокс.

Заметим, что тот же метод рассуждения можно применить к случаю, когда никакого движения нет, но среда за решеткой — кристаллическая (оптически анизотропная). Общее со случаем движущейся решетки заключается здесь в том, что в обоих случаях скорость света зависит от направления распространения. При этом в результате рассуждений, подобных только что приведенным, мы получим неверный результат.

Почему же рассуждение, приводящее к правильному результату в случае покоящейся решетки, в случае движущейся (а также в случае анизотропной среды) приводит к неправильным результатам? Ответ в том, что интерпретация решения (38), полученного для бесконечной решетки, здесь не соответствует интересующему нас опыту с конечной решеткой.

Если решетка бесконечная, то всюду, в каждой точке пространства позади нее, действует вся совокупность волн, входящих в правую часть выражения (38). Реальное значение имеет только эта их суперпозиция, их сумма, только ее можно наблюдать. Никаких максимумов и минимумов, соответствующих обычной картине спектров разных порядков, за бесконечной решеткой мы не получим. Интересующий же нас опыт состоит именно в том, что наблюдаются дифракционные максимумы, соответствующие отдельным членам суммы (38). Поэтому для выяснения вопроса нужно разобрать опыт, в котором можно отделить друг от друга члены суммы, входящей в (38).

С бесконечной решеткой без применения линз такой опыт осуществить нельзя. Решение же задачи с бесконечной решеткой, хотя для нее и учитывается правильная зависимость волнового числа от направления, т. е.

$$k = \frac{2\pi}{\lambda_0}[1 - \beta(\cos\psi - 1)], \tag{39}$$

не дает никакого ответа относительно результата такого опыта, так как постановка вопроса не соответствует условиям этого опыта. В частном случае покоящейся решетки можно было из решения задачи о бесконечной решетке получить ответ о результате интересующего нас опыта только потому, что в этом случае можно было дать отдельным членам суммы правильную интерпретацию. Обоснование же этой интерпретации (которая в таком виде, как для покоящейся решетки, для движущейся непригодна) дается, по существу, только решением задачи о дифракции на конечной решетке. Таким образом, необходимо решить задачу о дифракции на конечной решетке.

Решение задачи о дифракции плоской волны на ограниченной решетке получается по методу Релея в виде интеграла Фурье

$$E = \int g(u) e^{-i\left(ux+\sqrt{k^2-u^2}z\right)} du. \tag{40}$$

При этом для решетки, состоящей из m щелей, при нормальном падении волны на нее

$$g(u) = \text{const} \frac{\sin au}{au} \frac{\sin mDu}{\sin Du}, \tag{40'}$$

где $2D = d$ — постоянная решетки, $2a$ — ширина каждой щели. Найдем дифракционную картину на очень большом расстоянии от решетки (при $z \to \infty$); как можно показать, она подобна той дифракционной картине, которая получится при употреблении линзы в фокальной плоскости последней. Нужно продискутировать и вычислить интеграл (40) для больших значений расстояния $r = \sqrt{x^2 + z^2}$, предполагая при этом, что k зависит от угла ψ согласно (39): $k = k_0 [1 + \beta(1 - \cos\psi)]$, причем ψ связано с переменной u соотношением $k \sin\psi = u$. Это вычисление, по существу, необходимо при всякой дифракционной задаче (и в случае неподвижной решетки тоже), но в нашем случае необходимость этого расчета выявляется особенно резко. Нужно, таким образом, найти асимптотическое значение интеграла (40) для $r \to \infty$.

Для этого мы воспользуемся чисто математической теоремой, касающейся асимптотического выражения интеграла вида

$$J = \int_a^b g(u) e^{-ir\Phi(u)} du \tag{41}$$

при $r \to \infty$. Теорема эта состоит в том, что если в промежутке интегрирования производная $\Phi'(u)$ обращается в нуль, то при

$r \to \infty$ величина J убывает, как $\frac{1}{\sqrt{r}}$, и для J имеет место приближенное выражение

$$J \approx \sqrt{\frac{2\pi}{\pm \Phi''(u_0)\, r}}\, g(u_0)\, e^{-i\left[r\Phi(u_0) \mp \frac{\pi}{4}\right]}. \qquad (42)$$

Здесь u_0 — значение, для которого $\Phi'(u_0) = 0$. Предполагается при этом, что в промежутке интегрирования лежит только один корень этого уравнения. В формуле (42) нужно брать те или иные знаки в зависимости от знака $\Phi''(u_0)$: если $\Phi''(u_0) > 0$, то верхний знак, если же $\Phi''(u_0) < 0$ — нижний.

Для пояснения этой теоремы заметим, что если $g(u)$ имеет в интервале интегрирования конечное число максимумов и минимумов и непрерывна, то

$$\int g(u) \cos ru \cdot du$$

убывает при $r \to \infty$, как $1/r$, в чем можно убедиться путем интегрирования по частям. Причину этого легко понять, если учесть, что $\cos ru$ тем чаще меняет знак, чем больше r. Поэтому, чем больше r, тем полнее компенсируют друг друга положительные и отрицательные части интеграла. В результате величина интеграла оказывается порядка этого интеграла, взятого за один период $\cos ru$ (равный $1/r$), и имеет поэтому порядок величины $1/r$.

Для интеграла

$$\int g(u) \cos [r\Phi(u)]\, du$$

имеет место то же самое, за исключением тех точек, где $\Phi(u)$ меняется очень медленно, т. е. за исключением окрестности точки u_0, для которой $\Phi'(u_0) = 0$. Дальнейшее исследование интеграла по этой части промежутка интегрирования и приводит к формуле (42).

Применим теперь эту формулу к интегралу, входящему в (40). Мы получим решение, которое даст нам распределение дифракционного поля на очень больших расстояниях от ограниченной решетки. Это решение даст нам ответ на поставленную выше физическую задачу.

Заметим, что асимптотическая формула (42) применима к (40) при условии $r\lambda \gg m^2D^2$. Применяя (42) к (40), мы полагаем $x = r \sin \varphi$, $z = r \cos \varphi$, где φ — полярный угол точки наблюдения.

Тогда выражение (40) можно написать так:

$$E = \int g(u)\, e^{-ir\Phi} du, \qquad (43)$$

где

$$\left.\begin{aligned} &\Phi = k(\psi)\cos(\psi-\varphi),\\ &k \equiv k(\psi) = k_0[1+\beta(1-\cos\psi)],\\ &u = k(\psi)\sin\psi. \end{aligned}\right\} \quad (44)$$

Чтобы воспользоваться формулой (42), находим $\Phi'(u)$:

$$\Phi'(u) = \{-k(\psi)\sin(\psi-\varphi) + k'(\psi)\cos(\psi-\varphi)\}\frac{d\psi}{du},$$

где $k'(\psi) = \frac{dk}{d\psi} = k_0\beta\sin\psi$. Условие $\Phi'(u_0) = 0$ дает уравнение для нахождения соответствующего значения ψ_0

$$k(\psi_0)\sin(\psi_0-\varphi) - k'(\psi_0)\cos(\psi_0-\varphi) = 0. \quad (44')$$

Удерживая при вычислении только члены первого порядка относительно β и учитывая, что $k'(\psi_0)$ пропорционально β, получим отсюда

$$\psi_0 - \varphi = \frac{k'(\varphi)}{k_0} = \beta\sin\varphi,$$

$$u_0 = k(\varphi)\sin\varphi + k'(\varphi)\cos\varphi = k_0(1+\beta)\sin\varphi. \quad (45)$$

Далее, с той же точностью

$$\Phi(u_0) = k(\varphi).$$

Значение $\Phi''(u_0)$, входящее в амплитуду, достаточно вычислить, пренебрегая и первыми степенями β; оно равно тогда

$$\Phi''(u_0) = \frac{1}{k^2\cos\varphi^2} > 0.$$

Подставляя найденные значения в формулу (42), найдем

$$E = \left(\frac{2\pi}{r}\right)^{1/2} k\cos\varphi\cdot g(u_0)e^{-i\left[k(\varphi)r - \frac{\pi}{4}\right]},$$

причем u_0 определяется выражением (45).

Для решетки, согласно (40′),

$$g(u_0) = \text{const}\,\frac{\sin au_0}{au_0}\,\frac{\sin\frac{mdu_0}{2}}{\sin\frac{du_0}{2}}$$

имеет максимум при значении аргумента

$$u_0 = \frac{2\pi}{d} n,$$

где n — целое. Таким образом, максимумы поля получаются в направлениях, удовлетворяющих условиям

$$u_0 = k(1+\beta)\sin\varphi_{max} = \frac{2\pi n}{d},$$

т. е.

$$\sin\varphi_{max} = \frac{\lambda_0 n}{d(1+\beta)} = \frac{\lambda n}{d}.$$

Смещение Допплера будет, следовательно, обнаружено в полном согласии с принципом относительности [1].

При помощи только что изложенных соображений можно решить также один вопрос, касающийся аберрации. Пусть относи-

---

[1] [Полученный результат связан с тем, что при движении наблюдателя среда является оптически анизотропной, скорость света в ней зависит от направления. В оптически же анизотропной среде, как известно из кристаллооптики, направление луча не совпадает с направлением нормали к волне. Направление от решетки на дифракционный максимум, определяемое углом φ_{max}, есть направление луча, соответствующего волне, направление нормали к которой, определяемое углом ψ_{max}, удовлетворяет условию дифракции.

$$k(\psi_{max})\cdot\sin\psi_{max} = \frac{2\pi n}{d}.$$

Легко убедиться в том, что вычисления, которые были проделаны для нахождения угла φ_{max} при применении асимптотической формулы, совпадают с теми, которые нужно проделать для нахождения луча, соответствующего данной волне, по правилам, обычно приводимым в кристаллооптике.

Действительно, для нахождения направления луча, соответствующего волне направления ψ, нужно, как известно, построить огибающую семейства плоских волн разных направлений и одинаковой фазы (считая, что в начале координат их фазы одинаковы). Направление на точку пересечения двух соседних волн с направлениями ψ и $\psi + d\psi$ и даст направление луча.

Уравнение этого семейства плоских волн в полярных координатах r, φ, очевидно, можно написать так:

$$r\mathbf{\Phi} = \text{const},$$

где $\mathbf{\Phi} = \mathbf{\Phi}(\varphi, \psi)$ имеет то же значение, что и в тексте. Угол ψ — параметр семейства. Направление φ на точку пересечения двух соседних волн с направлениями ψ и $\psi + d\psi$ найдется из уравнения

$$\frac{\partial\mathbf{\Phi}(\varphi, \psi)}{\partial\psi} = 0,$$

совпадающего с (44′).]

тельно эфира движется наблюдатель, измеряющий направление света, падающего от звезды. Для простоты заменим трубу, с помощью которой наблюдается звезда, щелью. Скорость света по отношению к наблюдателю равна $c = c_0\,(1 + \beta \cos\psi)$. Предполагая, что наблюдатель движется перпендикулярно к лучу света, и проводя наши вычисления для угла наблюдения (угла аберрации) ψ, можно получить хорошо известное значение

$$\psi \approx \sin\psi = \beta = \frac{v}{c_0}\,.$$

Вопрос возникает в том случае, если допустить в эфире наличие дисперсии. Известно, что все обычные методы измерения скорости света в среде с дисперсией дают *групповую* скорость. Какая скорость c_0 входит в формулу для угла аберрации при наличии дисперсии в эфире? Релей думал, что путем измерения угла аберрации мы, в отличие от других методов, определяем *фазовую* скорость. Однако Эренфест показал, что и в выражении для угла аберрации в диспергирующей среде c_0 — групповая скорость. Этот результат Эренфеста легко получить нашим методом, если учесть, что при наличии дисперсии волны разных направлений будут иметь не только разную длину волны, но и разные частоты[1].

ВОСЬМАЯ ЛЕКЦИЯ

(7. IV 1933 г.)

Опыт, предложенный Эйнштейном, для решения вопроса о волновой или квантовой природе света. О реальности разложения колебаний на синусоидальные. Действие импульса на резонатор. Связь между временем затягивания спектрального аппарата и его разрешающей способностью

Для решения вопроса о том, что представляет собой свет — волны или кванты — Эйнштейн предложил сделать следующий опыт[2]. Мимо щели в экране пролетает частица («каналовая»), излучающая монохроматический свет. Излученный ею и прошед-

---

[1] *П. Эренфест.* Ann. Physik, 33, 1571, 1910; *Релей.* Phil. Mag., 22, 130, 1919 (Sci. Pap., т. VI, стр. 42); *М. П. Свешникова.* ЖРФХО, 59, 377, 1927.
[2] *А. Эйнштейн.* Naturwiss., № 14, 300, 1926; Berl. Ber., 334, 1926.

ший через щель свет наблюдается спектральным аппаратом (см. рис. 3).

С волновой точки зрения за щелью свет от такой частицы представляет собой оборванный волновой цуг, отрезок синусоиды, т. е. он не будет монохроматичным и будет тем менее монохроматичным, чем меньше время пролета частицы мимо щели. Спектральная линия, излученная частицей, должна оказаться уширенной при наблюдении за щелью.

С точки зрения представления о световых квантах (фотонах) расширения линии не должно быть, так как каждый квант монохроматичен, он может изменить частоту, только изменив свою энергию. При прохождении через щель квант не теряет энергии, значит, за щелью будет монохроматический свет.

Бор указал на то, что и с квантовой точки зрения уширение должно иметь место [1]. Однако мне кажется, что вся постановка вопроса лишена смысла.

Нужно выяснить прежде всего, что значит «монохроматический квант», действие которого ограничено во времени. В волновой теории под монохроматическим светом понимают световые колебания, имеющие вид неограниченной во времени синусоиды. Таким образом, ограниченность во времени и монохроматичность исключают друг друга. В теории световых квант нельзя связать монохроматичность с синусообразностью, поэтому нужно дать какое-то другое определение монохроматичности. Определение монохроматичности, даваемое в волновой теории, содержит в себе возможность экспериментального контроля монохроматичности колебаний, ибо синусоидальная волна дает в спектральном

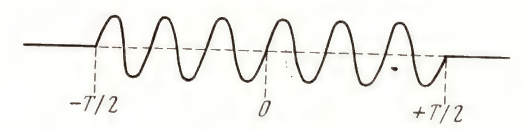


Рис. 16

приборе, например приборе с дифракционной решеткой, резкую линию. Это показывает теория спектральных приборов, основанная на волновых представлениях.

Давая определение монохроматичности в теории световых квант, также необходимо связать это определение с теорией действия спектральных приборов, но основанной на квантовых пред-

[1] [В своих лекциях Л. И. Мандельштам ссылался на опыт, опубликованный Руппом, результаты которого подтверждали наличие уширения. Однако потом работы Руппа оказались дискредитированными, так что теперь ссылка на них не имеет значения.]

ставлениях. Только указание на характер действия аппаратов в рамках данной теории придает физический смысл всем ее утверждениям.

Утверждение Эйнштейна о том, что квант после прохождения через щель даст резкую линию, неизбежно уже содержит молчаливое допущение в отношении механизма действия воспринимающего аппарата.

Рассмотрим вопрос о монохроматичности более подробно, прежде всего с волновой точки зрения. Выясним, как реагирует спектральный аппарат на ограниченный волновой цуг. Колебания для такого ограниченного импульса продолжительности T изображаются функцией (рис. 16).

$$\left.\begin{aligned} &f(t)=\sin nt, && -\frac{T}{2}<t<+\frac{T}{2},\\ &f(t)=0, && |t|>\frac{T}{2}. \end{aligned}\right\} \qquad (46)$$

При этом мы будем считать, что $nT = 2\pi N$, где N — целое число, чем мы обеспечим непрерывность функции $f(t)$ при $t = \pm\frac{T}{2}$; $f(t)$ — нечетная функция. Ее разложение в интеграл Фурье таково:

$$\left.\begin{aligned} f(t)&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}G(u)\sin ut\,du,\\ G(u)&=\sqrt{\frac{2}{\pi}}\int_{-\infty}^{+\infty}f(t)\sin ut\,dt. \end{aligned}\right\} \qquad (47)$$

Разложение $f(t)$ в интеграл Фурье можно написать и в комплексной форме

$$\left.\begin{aligned} f(t)&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}g(u)\,e^{iut}\,du,\\ g(u)&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}f(t)\,e^{-iut}dt=\begin{cases}\dfrac{1}{2i}G(u) & \text{при } u>0,\\ -\dfrac{1}{2i}G(-u) & \text{при } u<0.\end{cases} \end{aligned}\right\} \qquad (48)$$

Раньше мы всегда пользовались комплексной формой (48), здесь же естественной является действительная форма (47). Это связано с различием в понятии монохроматичности в отношении пространственной зависимости для волн, с одной стороны, и моно-

хроматичности в отношении зависимости колебаний от времени — с другой, на которое уже указывалось (лекция 6, стр. 50). Для всех задач, касающихся распространения в пространстве волн определенной частоты, естественной является форма, дающая в зависимости от пространственной координаты x (мы говорим о волнах в пространстве одного измерения) разложение на волны типа e^{iux}. Для всех же вопросов, касающихся изменения во времени, естественным является разложение Фурье на действительные синусоиды.

В обоих случаях наблюдаемые величины действительны, и физический смысл имеют действительные числа, а при применении комплексных величин — их действительные (или мнимые) части. Однако в первом случае мы имеем уравнение в частных производных и рассматриваем его решения вида $e^{int}\varphi(x)$. При этом решения, соответствующие $\varphi = e^{iux}$ и $\varphi = e^{-iux}$, имеют разный смысл — это волны, идущие в противоположных направлениях [их действительные части $\cos(nt + ux)$ и $\cos(nt - ux)$, а мнимые $\sin(nt + ux)$ и $\sin(nt - ux)$]. В случае же зависимости колебания от времени e^{int} и e^{-int} изображают колебания, совпадающие или отличающиеся только фазами (их действительные части совпадают, мнимые отличаются только знаком). Таким образом, для первого случая естественным элементом разложения является e^{iux} ($u \lessgtr 0$) и соответственно комплексная форма интеграла Фурье, для второго же случая элементом является совокупность $\sin nt$ и $\cos nt$ и соответственно действительная форма интеграла. Конечно, и при разложении в зависимости от времени можно пользоваться комплексной формой разложения — математически это иногда удобно, и мы будем это делать,— но здесь эта форма разложения не имеет того значения, как в случае пространственной зависимости.

Простейший спектральный аппарат — это резонатор, у которого можно менять собственную частоту. Примером такого резонатора в радиотехнике является обычный волномер. В оптике в качестве такого аппарата можно применять пары металла, поглощающие свет в области резонансной линии, например пары натрия; частоту поглощающей линии можно при этом изменять, помещая пар в магнитное поле (используя явление Зеемана). Кроме резонатора, спектральным аппаратом является и дифракционная решетка. Действие всех этих аппаратов подчиняется линейным уравнениям, поэтому действие на них любого импульса можно представить как действие составляющих его синусоид. Часто ставят вопрос о «реальности» разложения колебания на синусоидальные составляющие. Что реально в равенстве (47), левая часть или же правая — совокупность синусоид? В такой форме вопрос лишен смысла, так как равенство (47) — это чисто

математическое утверждение, т. е. справа и слева стоит одно и то же.

Вопрос приобретает смысл, если его ставят в связи с аппаратами, воспринимающими колебания. В зависимости от свойств рассматриваемого аппарата бывает целесообразным пользоваться либо левой, либо правой частью равенства (47) (или заменяющего его разложения на синусоиды, например ряда Фурье). Поясним это простым примером. Пусть

$$f(t) = (1 + k \cos \omega t) \cos nt. \tag{49}$$

С помощью тригонометрического преобразования $f(t)$ можно представить в виде суперпозиции синусоид

$$f(t) = \cos nt + \frac{1}{2} \cos (n + \omega) t + \frac{1}{2} \cos (n - \omega) t. \tag{50}$$

Будем считать, что $\omega \ll n$. Можно сказать, во-первых, что [согласно (49)] $f(t)$ представляет собой синусообразное колебание с медленно изменяющейся амплитудой и, во-вторых, что [согласно (50)] это суперпозиция синусоидальных колебаний с тремя разными частотами. Очевидно, что и то и другое тождественно. Однако как только мы принимаем во внимание воспринимающий аппарат, то два представления уже не эквивалентны. Опыт с камертоном, звук которого во время распространения по воздуху периодически (с периодом $2\pi/\omega$) прерывается рукой, опускаемой между камертоном и воспринимающим аппаратом, иллюстрирует это. Звуковые колебания, приходящие к аппарату, можно при этом приближенно считать имеющими вид только что рассмотренной функции $f(t)$. Если воспринимающим аппаратом является ухо, то правильно первое представление, так как ухо не разделяет трех тонов (50) и слышит один тон переменной громкости. Если же воспринимающим аппаратом является другой камертон, то правильно второе представление, так как этот второй камертон откликается только в том случае, когда он настроен на одну из трех частот: n, $n + \omega$, $n - \omega$.

Вернемся теперь к вопросу о действии ограниченного волнового цуга (46) на резонатор. Разлагая эту функцию $f(t)$ в интеграл Фурье в форме (48), получим для нее

$$g(u) = \sqrt{\frac{2}{\pi}}\, n \frac{\sin \frac{(u-n)T}{2}}{u^2 - n^2},$$

так что

$$f(t) = \frac{n}{\pi} \int\limits_{-\infty}^{+\infty} \frac{\sin \frac{(u-n)T}{2}}{u^2 - n^2} e^{iut}\, du. \tag{51}$$

Это наложение бесконечных синусоид, тянущихся от $t = -\infty$ до $t = +\infty$. Но до момента $t = -T/2$ функция $f(t)$ равна нулю. Поэтому до момента $t = -T/2$ колебание не будет действовать ни на какой аппарат. Из этого делают иногда вывод, что синусоиды «нереальны». Однако, учитывая эквивалентность левой и правой частей выражения (51), можно заключить лишь то, что действия всех бесконечных синусоид, на которые разлагается $f(t)$, складываются так, что результатом действия всей их совокупности является нуль для всех моментов $t < -T/2$. Мы докажем сейчас, что для резонатора со сколь угодно малым затуханием это действительно выполняется.

Уравнение колебаний резонатора под действием импульса $f(t)$ имеет вид

$$y + 2\delta y + \omega_0^2 y = f(t), \qquad (52)$$

где δ — постоянная затухания. Если на резонатор действует сила e^{iut}, то решение будет

$$y_1 = \frac{e^{iut}}{\omega_0^2 - u^2 + 2iu\delta}\,. \qquad (53)$$

Если же на резонатор действует сила

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(u)\, e^{iut}\, du,$$

то решение уравнения (52) получится из (53) путем умножения последнего выражения на $\frac{g(u)\,du}{\sqrt{2\pi}}$ и интегрирования

$$y(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{g(u)\, e^{iut}\, du}{\omega_0^2 - u^2 + 2iu\delta}\,.$$

Подставим сюда значение (48) для $g(u)$ и, заменяя порядок интегрирования, получим

$$y(t) = \frac{1}{2\pi} \int f(\xi)\, d\xi \int_{-\infty}^{+\infty} \frac{e^{i(t-\xi)u}\, du}{\omega^2 - u^2 + 2iu\delta}\,.$$

Это выражение должно быть нулем для $t < -T/2$. Однако для $t > T/2$, т. е. для моментов времени после окончания действия силы, y может и не быть нулем, резонатор будет продолжать колебаться. При доказательстве должен, таким образом, играть существенную роль знак t. Доказательство должно быть «одно-

боким» — оно должно давать нуль для $t < -T/2$ и может его не давать для $t > T/2$. Для вычисления воспользуемся методом интегрирования по комплексной плоскости и рассмотрим интегрирование по u:

$$\int_{-\infty}^{+\infty} \frac{e^{i(t-\xi)u}\, du}{\omega_0^2 - u^2 + 2iu\delta}.$$

Будем считать, что $t < -T/2$. Достаточно рассмотреть значения интеграла для значений $\xi > -T/2$, так как при $\xi < -T/2$ $f(\xi) = 0$. Значит, для этих моментов времени $t, t-\xi < 0$. Рассматривая u как комплексную переменную, мы имеем сначала путь интегрирования по действительной оси от $-\infty$ до $+\infty$. Путь интегрирования можно деформировать и, если в нижней полуплоскости нет полюсов подынтегрального выражения, замкнуть его по бесконечно большому полукругу в нижней полуплоскости (рис. 17). Полюсами подынтегрального выражения будут корни уравнения

$$u^2 - 2i\delta u - \omega_0^2 = 0,$$

которые лежат в верхней, а не в нижней полуплоскости. Таким образом, интересующий нас интеграл равен интегралу по пути, охватывающему нижнюю полуплоскость и, следовательно, равен нулю, так как $t - \xi < 0$ и в этой полуплоскости $\operatorname{Im} u < 0$. Таким образом, $y = 0$ при $t < -T/2$.

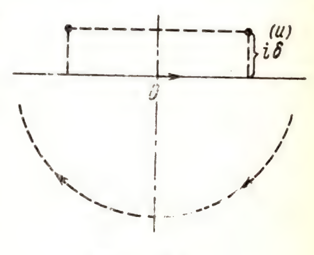


Рис. 17

При этом доказательстве мы нигде не пользовались специальным видом функции $f(t)$, а учитывали лишь то, что она равна нулю для $t < -T/2$. Тот же результат легко получить тем же способом и для любой сколь угодно сложной линейной колебательной системы, используемой в качестве резонатора, если только она работает с потреблением энергии (ее свободные колебания — затухающие). В этом случае полюсы соответствующего выражения будут лежать в верхней полуплоскости и доказательство останется в силе.

Вернемся к решению уравнения для резонатора, на который действует оборванная с обеих сторон синусоида. Пользуясь (51) и (53), получим

$$y = \frac{n}{\pi} \int_{-\infty}^{+\infty} \frac{\sin\frac{(u-n)T}{2}\, e^{iut}\, du}{(u^2 - n^2)(\omega_0^2 - u^2 + 2i\delta u)}. \tag{54}$$

Заметим, что этот способ решения задачи в известном смысле специален. Здесь как будто бы не вошли начальные условия, без задания которых решение дифференциального уравнения неопределенно. Мы рассматривали задачу для t от $-\infty$ до $+\infty$ и тем самым учли при решении все воздействия за этот промежуток времени. Легко убедиться, однако, что это решение соответствует начальным условиям $y\left(-\frac{T}{2}\right)=0$, $\dot{y}\left(-\frac{T}{2}\right)=0$. Можно также сказать, что это решение соответствует любым начальным условиям для $t=-\infty$.

Большинство приборов в оптике определяет величину $J = \int\limits_{-\infty}^{+\infty} |y|^2\,dt$. Эта величина грубо приближенно определяет, например, почернение фотопластинки. Если, скажем, я фотографирую резонансное свечение ртутного или натриевого пара под действием монохроматического излучения, действующего на него в течение некоторого времени, то почернение будет определяться именно этим интегралом, распространенным до $+\infty$, так как нужно учесть и затухающее излучение резонаторов — атомов пара — после прекращения освещения. Для вычисления величины J воспользуемся теоремой Релея, согласно которой

$$\int\limits_{-\infty}^{+\infty} |f|^2\,dt = \int\limits_{-\infty}^{+\infty} |g|^2\,du,$$

где g и f — функции, сопряженные по Фурье [согласно (48)]. Тогда, учитывая (54), в нашем случае действия ограниченной синусоиды на резонатор, для величины J получим

$$J = \int\limits_{-\infty}^{+\infty} |y|^2\,dt = \int\limits_{-\infty}^{+\infty} |g|^2\,du = \frac{2n^2}{\pi}\int\limits_{-\infty}^{+\infty} \frac{\sin^2\frac{(u-n)T}{2}\,du}{(u^2-n^2)^2\,[(\omega_0^2-u^2)^2+4\delta^2u^2]}\,.$$

Это выражение позволяет убедиться в том, что значение J существенно зависит от длительности сигнала T по сравнению с временем затухания резонатора $1/\delta$.

Рассмотрим два крайних случая: 1) когда $nT \gg \omega_0/\delta$ и 2) когда $nT \ll \omega_0/\delta$.

В первом из них длительность импульса T очень велика по сравнению с временем затухания $1/\delta$ (n и ω_0 можно считать величинами одного порядка), во втором имеет место обратное соотношение.

В первом случае, когда $nT \gg \omega_0/\delta$ или (что практически то же самое) когда $T \gg 1/\delta$, фактор $[(\omega_0^2-u^2)^2+4\delta^2u^2]^{-1}$ мало

меняется в интервале значений u ширины порядка $1/T$ вокруг значения $u = n$, в котором $\frac{\sin^2 \frac{(u-n)T}{2}}{(u^2-n^2)^2}$ имеет заметную величину. Поэтому этот фактор можно вынести за знак интеграла, положив в нем $u = n$, так что

$$J = \frac{2n^2}{\pi} \frac{1}{(\omega_0^2 - n^2)^2 + 4\delta^2 n^2} \int_{-\infty}^{+\infty} \frac{\sin \frac{(u-n)T}{2}}{(u^2-n^2)^2} du = \text{const} \frac{T}{(\omega_0^2 - n^2)^2 + 4\delta^2 n^2}.$$

Таким образом, на *распределение* величины J в спектре, т. е. на зависимость J от ω_0, не влияет длительность импульса, и она определяется только свойствами спектрального прибора — затуханием δ нашего резонатора. Распределение J по спектру такое же, как и в случае действия на наш спектральный аппарат бесконечной синусоиды.

В этом случае мы не можем поэтому сделать из спектрального распределения никакого заключения о длительности импульса.

Во втором случае продолжительность импульса мала по сравнению со временем затухания резонатора: $nT \ll \omega_0/\delta$ ($T \ll 1/\delta$). Здесь, наоборот, медленно меняется множитель $\frac{\sin \frac{(u-n)T}{2}}{(u^2-n^2)^2}$. Поэтому мы можем вынести его из-под знака интеграла, положив в нем значение u равным ω_0, для которого второй множитель $[(\omega_0^2 - u^2)^2 + 4\delta^2 u^2]^{-1}$ имеет максимум.

Тогда

$$J = \frac{2n^2}{\pi} \frac{\sin^2 \frac{(\omega_0 - n)T}{2}}{(\omega_0^2 - n^2)^2} \int_{-\infty}^{+\infty} \frac{du}{\omega_0^2 - u^2 + 4\delta^2 u^2} = \text{const} \frac{\sin^2 \frac{(\omega_0 - n)T}{2}}{(\omega_0^2 - n^2)^2}.$$

Итак, в этом случае J оказывается пропорциональным квадрату коэффициента Фурье функции $f(t)$, действующей на резонатор. Распределение J в спектре определяется теперь не свойствами спектрального аппарата, а длительностью действующего на аппарат импульса и его формой. Таким аппаратом можно изучить свойства импульса, так как он дает распределение спектра импульса. Чем меньше будет длительность импульса T, тем расплывчатей получится спектральное распределение в аппарате; ширина этого распределения Δn порядка $1/T$.

В промежуточном случае распределение в спектре аппарата будет определяться и свойствами импульса, и свойствами аппарата.

Таким образом, разрешающая способность резонатора как спектрального аппарата определяется временем его затухания $T_1 = 1/\delta$. Это его «временная постоянная». Условие разрешения для данного прибора можно записать в виде $T_1 \gtrsim T$, где T — по-прежнему длительность исследуемого импульса, или же вводя разность частот в спектре, которую надо разрешить,—в виде $\Delta n \cdot T_1 \gtrsim 1$. Но чем больше время затухания резонатора T_1, тем больше нужно времени, чтобы установилась амплитуда вызванных исследуемым импульсом колебаний, тем больше время, требуемое для измерения с таким аппаратом.

Итак, мы приходим к выводу, что *время, нужное для того, чтобы спектральным аппаратом — резонатором — произвести измерение, тем больше, чем больше его разрешающая сила — острота настройки.* Другими словами, чем точнее нужно определить спектр импульса, тем более длительным должен быть процесс измерения.

Это положение, показанное нами для резонатора, справедливо для *любого* спектрального прибора. Мы докажем его еще для дифракционной решетки.

ДЕВЯТАЯ ЛЕКЦИЯ

(25.IV 1933 г.)

Разрешающая способность и время затягивания дифракционной решетки. Спектральные приборы создают периодичность. Доказательство того, что предложенный Эйнштейном опыт не есть experimentum crucis. Объяснение дифракции в рамках теории световых квантов

Мы рассмотрели действие резонатора как спектрального аппарата. Рассмотрим теперь другой спектральный аппарат — дифракционную решетку. Как известно, разрешающая способность решетки определяется числом периодов в ней. Если a — длина решетки, d — ее период, то $m = \frac{a}{d}$. Разрешение двух частот, n и $n + \Delta n$, в спектре первого порядка возможно, если $\frac{n}{\Delta n} < m$. В применении к анализу спектра оборванной синусоиды длины T это приводит к следующему условию, при котором можно по уши-

рению в спектре обнаружить обрыв синусоиды:

$$nT < m.$$

Понятие о времени затухания спектрального прибора T_1 может быть перенесено и на решетку, у которой, так же как и у резонатора, есть определенная «временная постоянная». Действительно, за решеткой импульс длится то же самое время, что и до нее, только в том случае, если мы наблюдаем в направлении, перпендикулярном к решетке. При наблюдении же под углом длительность сигнала, измененного действием решетки, будет больше, чем до нее, так как волны, приходящие от разных точек решетки, имеют разность хода. Время затягивания равно $T_1 = \frac{r}{c}$, где (рис. 18) r — разность хода для волн, идущих от краев решетки. Это и есть временная постоянная решетки. Учитывая, что для спектра первого порядка $\frac{r}{\lambda} = \frac{a}{d} = m$, а значит,

$$T_1 = \frac{\lambda m}{c} = 2\pi \frac{m}{n},$$

условие разрешения двух частот, n и $n + \Delta n$, можно записать так:

$$\Delta n \cdot T_1 \gtrsim 1,$$

а условие для возможности обнаружить влияние продолжительности сигнала на его спектр так:

$$T_1 \gtrsim 2\pi T.$$

Оба эти условия (отвлекаясь от численного множителя 2π, связанного с некоторым произволом в определении T_1) совпадают с таковыми для резонатора. Это общие условия, годные в таком виде для любого спектрального аппарата. Таким образом, связь между временем, нужным для измерения спектральным аппаратом и определяемым его временной постоянной, и его разрешающей способностью носит общий характер и относится к любому спектральному аппарату. Нельзя построить спектральный аппарат с малым временем затягивания и большой разрешающей способностью.

Эти общие соображения позволяют разрешить следующий вопрос. Пусть на дифракционную решетку падает сигнал, имеющий вид, изображенный на рис. 19: периодически (с периодом T) следующие друг за другом отрезки синусоиды частоты n, чередующиеся с перерывами в сигнале. Разлагая такую функцию в ряд Фурье, получим совокупность частот $n \pm m\omega$, где $m = 0, 1, 2, \ldots$

При достаточной разрешающей способности решетки в спектральном приборе мы получим равноотстоящие линии. Это значит, что в тех местах, где получаются эти линии, колебания будут незатухающими синусоидами. Но ведь в промежутках времени между отрезками синусоид на прибор не попадало никакого излучения! Теперь этот парадокс разрешается легко: дело в том, что решетка реагирует на сигнал не синхронно, она затягивает сигналы, перекрывая благодаря этому «зоны молчания» падающего излучения.

Вообще периодичность, имеющаяся в отдельных монохроматических линиях спектра импульса, разрешенного решеткой или другим спектральным аппаратом, создана самим этим аппаратом. Пусть, например, на решетку падает импульс, имеющий вид отдельного короткого толчка. При действии решетки получаются цвета, т. е. периодические колебания. Решетка повторит сигнал m раз — по числу ее штрихов, и, например, красный цвет мы увидим в том месте, где период между каждым повторением импульса равен периоду красного света.

Остановимся теперь еще на одном общем вопросе, связанном с распространением несинусоидальных импульсов на плоскости (в двух измерениях) и в пространстве трех измерений. Мы рассматривали до сих пор дифракционные задачи в основном для двух измерений. Такое упрощение было вполне допустимо, потому что для синусообразных колебаний разница между плоским и пространственным случаем несущественна. Теперь мы перешли к распространению импульсов. Для распространения немонохроматических импульсов различия, которые получаются при двух и при

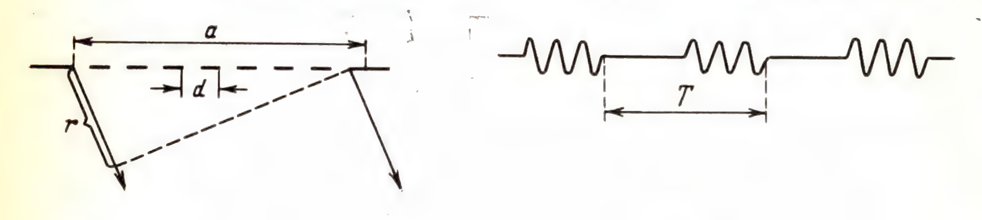


Рис. 18 Рис. 19

трех измерениях, оказываются более существенными. Вообще распространение импульса в пространстве нечетного числа измерений (1, 3...) имеет несколько иной характер, чем в пространстве четного числа измерений (2, 4, ...). Для нечетного числа измерений у распространяющегося импульса есть не только резкое начало, но и резкий конец. Для четного же числа измерений есть резкое начало импульса, но нет резкого конца.

Чтобы выяснить этот вопрос, поставим его несколько иначе. Пусть мы имеем волновое уравнение в трех измерениях

$$\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2}. \quad (55)$$

Если φ не зависит от y, то получаем волновое уравнение в двух измерениях

$$\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial z^2}, \quad (56)$$

к которому мы пришли бы сразу, если бы рассматривали, например, колебания мембраны.

В трехмерном случае дело обстоит так: пусть в начальный момент в безграничной трехмерной среде задано возмущение

$$\varphi(x, y, z, 0) = f(x, y, z), \qquad \dot{\varphi}(x, y, z, 0) = F(x, y, z).$$

Мы можем тогда найти возмущение $\varphi(x, y, z, t)$ для любого момента времени t. Решение этой задачи — интегрирование волнового уравнения (55) в трех измерениях при заданных начальных условиях — было дано Пуассоном [1]. Оно дает определенную картину распространения возмущения и сводится к следующему. Пусть первоначально возмущение ограничено определенной конечной областью пространства G. Вокруг точки P (с координатами x, y, z), для которой мы хотим определить $\varphi(x, y, z, t)$, надо описать сферу S_t, радиус которой r равен ct; он растет со временем со скоростью c. До момента $t_1 = \frac{r_1}{c}$ (см. рис. 20), в который эта сфера S_t коснется области G, поле φ в точке P будет нулем. Оно отлично от нуля только в те моменты времени t, когда сфера S_t пересекает область G (конечно, в отдельные моменты оно может при этом обращаться в нуль). После же момента $t_2 = \frac{r_2}{c}$, когда вся область первоначального возмущения G окажется внутри сферы S_t, поле в точке P опять будет равно нулю.

Таким образом, возмущение в P начинается в момент $t_1 = \frac{r_1}{c}$ и кончается в момент $t_2 = \frac{r_2}{c}$ (где r_1 и r_2 — наименьшее и наибольшее расстояния от точки P до точек области G). Значит, при распространении импульса в трех измерениях у такой волны есть резко выраженный передний фронт и резко выраженный обрыв в конце ее.

Иначе обстоит дело в плоском (двухмерном) случае, когда φ — решение уравнения (56). Чтобы свести решение задачи в этом слу-

[1] *Релей*. Теория звука, т. II, § 273.

чае на только что рассмотренное решение для трехмерного случая, рассуждаем так. Будем считать, что в трехмерном случае начальное возмущение не зависит от y

$$\varphi(x, y, z, 0) = f(x, z), \quad \dot{\varphi}(x, y, z, 0) = F(x, z).$$

Тогда, очевидно, и для всего дальнейшего времени поле φ не будет зависеть от y и φ будет удовлетворять двухмерному уравнению (56). В трехмерном пространстве это значит, что первоначальное возмущение заполняло бесконечный цилиндр G с направляющими по оси y. Чтобы решить, что будет происходить в какой-нибудь точке P, можно в трехмерном пространстве провести приведенное построение со сферой S_t. Пока сфера S_t не коснется ближайшей к P точки образующей цилиндра G, поле φ в точке P будет равно нулю. Таким образом, резкий передний фронт существует и в этом случае. Но теперь никогда не будет такого момента, когда *весь* цилиндр G будет внутри сферы S_t. Значит, резкого обрыва, резкого конца импульса в этом случае не будет. У двухмерной волны есть, так сказать, хвост, который никогда

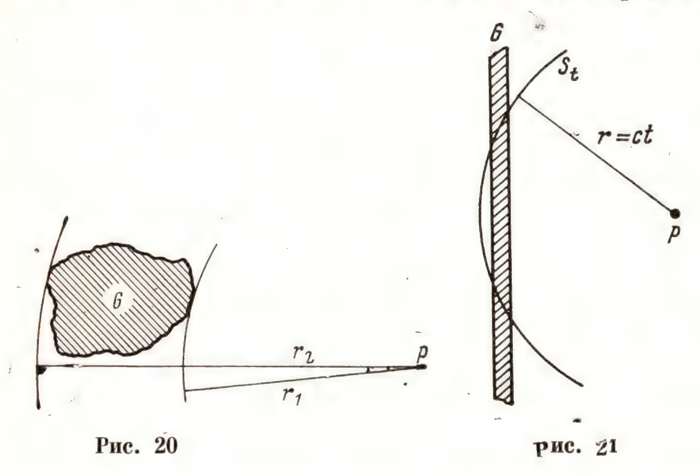


Рис. 20 Рис. 21

не пропадает. Этим распространение волн в двух измерениях резко отличается от распространения в трех измерениях.

Наше рассуждение наглядно, но может показаться неубедительным, так как тот же переход можно сделать от трех измерений и к одному измерению. Нужно только принять, что начальное возмущение, а следовательно и поле, для любого момента времени зависит только от одной координаты x. В этом одномерном случае вся область начального возмущения (имеющая вид слоя, ограниченного двумя плоскостями, перпендикулярными к оси x)

тоже никогда не будет находиться внутри сферы S_t. Между тем для одномерного случая резкий обрыв в конце импульса есть. Это объясняется тем, что после прохождения сигнала, когда сфера находится в положении, показанном на рис. 21, получается такая компенсация, что для всех этих моментов φ в точке P — нуль, и поэтому резкий обрыв получается. Для двух измерений такой компенсации не получается.

Во всяком случае, более точное исследование показывает, что при нечетном числе измерений пространства обрыв в конце импульса есть, а при нечетном числе измерений его нет [1].

Вернемся теперь к опыту, предложенному Эйнштейном (лекция 8, стр. 59). Посмотрим, в какой степени при постановке опыта можно соблюсти условие $T \ll T_1$, необходимое для надежного наблюдения уширения линии. Вместо щели можно взять дифракционную решетку. Ее период, соответствующий ширине щели, может быть порядка $d = 10^{-3}$ *см*. Скорость каналовых частиц, служащих источником света, порядка 10^{-8} *см/сек*. Тогда время пролета T будет порядка 10^{-11}*сек*. Если взять в качестве спектрального аппарата решетку размером 5 *см*, то $T_1 = \frac{5}{3 \cdot 10^{10}} \approx$ $\approx 2 \cdot 10^{-10}$. Следовательно, условие $T < T_1$ можно осуществить даже без применения интерференционных спектральных аппаратов и обнаружить уширение возможно.

Однако Бор показал, что и с точки зрения теории световых квантов должно получиться то же самое, что с волновой точки зрения, так что, в сущности, опыт, предложенный Эйнштейном, ничего не решает (нужно, конечно, иметь в виду, что сейчас [1933 г.] весь вопрос имеет лишь историческое значение; предложение Эйнштейна относится к 1926 г.).

Рассуждения Бора сводятся к следующему. Интересующий нас случай источника света, движущегося относительно щели и относительно наблюдателя со спектральным аппаратом, очевидно, эквивалентен (по принципу относительности) случаю, когда щель и наблюдатель движутся относительно покоящегося источника света. Мы уже рассматривали с точки зрения волновой теории (лекция 7) дифракцию на движущейся щели и видели, что получающаяся при этом картина отличается от картины дифракции на покоящейся щели только наличием явления Допплера для всех волн. Интересующий нас случай, когда и решетка и наблюдатель движутся, получится отсюда переходом к движущейся вместе со щелью системе координат, что вызовет допплеровское смещение частот всех волн. Таким образом, волновое объяснение

---

[1] *П. С. Эренфест.* Ann. Physik, 1917; см. также *Р. Курант, Д. Гильберт*, Методы математической физики. ГТТИ, 1945, т. II, гл. VI, § 9.

опыта Эйнштейна основано на наличии, во-первых, явления дифракции, и, во-вторых, явления Допплера. Но оба эти явления — и эффект Допплера и дифракция — объясняются с точки зрения теории световых квантов и при этом получаются те же формулы, что и с волновой точки зрения, а значит, говорит Бор, опыт Эйнштейна по теории световых квантов должен дать тот же результат, что и по волновой теории.

Объяснение явления дифракции с точки зрения теории световых квантов было дано в работе Эпштейна и Эренфеста и состояло в следующем [1]. Они вывели прежде всего соотношение для направления лучей, дифрагированных решеткой,

$$\sin\varphi - \sin\varphi_0 = \frac{m\lambda}{d} \tag{57}$$

путем применения к световым квантам и решетке законов сохранения энергии и импульса и основного соотношения старой боровской квантовой механики. Дифракция рассматривалась при этом как соударения светового кванта и решетки. До соударения энергия светового кванта равна $h\nu$, а слагающая импульса, параллельная плоскости решетки, есть

$$p_x^0 = \frac{h\nu}{c}\sin\varphi_0.$$

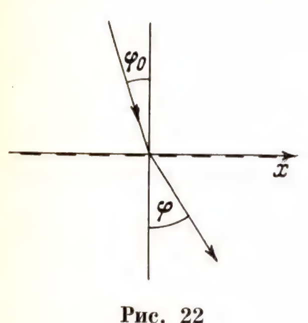


Рис. 22

Решетка до столкновения покоится, импульс ее — нуль. После столкновения x-компонента импульса кванта равна

$$p_x = \frac{h\nu}{c}\sin\varphi,$$

а компонента импульса решетки, получившей при ударе скорость с x-компонентой v, равна Mv, где M — масса решетки. Так как масса решетки M очень велика, то скорость ее v ничтожна. Поэтому энергией, переданной решетке, можно пренебречь и считать, что энергия кванта, а значит, и его частота при дифракции не изменяются. Закон сохранения импульса для x-компоненты его дает условие

$$p_x + Mv = p_x^0,$$

или

$$\frac{h\nu}{c}\sin\varphi + Mv = \frac{h\nu}{c}\sin\varphi_0. \tag{58}$$

---

[1] *P. Epstein, P. Ehrenfest.* Proc. Nat. Ac. USA, 10, 133, 1924.

После удара решетка движется по оси x с постоянной скоростью. Прохождение ее штрихов — периодический процесс. По старой квантовой механике для периодического движения должно выполняться условие

$$\oint P_x dx = mh,$$

где интеграл от импульса P_x системы, совершающей периодическое движение, нужно взять за период, m — целое положительное число. В нашем случае $P_x = Mv$, $\oint P_x dx = Mvd$, где d — период решетки. Поэтому $Mv = \frac{mh}{d}$. Подставляя это значение Mv в (58), получаем дифракционную формулу (57).

Эта теория, как и вся старая боровская квантовая теория, не может решить, какие целые значения m возможны для данной решетки, и не может дать распределения интенсивности по дифракционным спектрам разных порядков.

Вопросы, связанные с интенсивностями линий в спектре атома, разрешались в старой боровской теории с помощью «принципа соответствия». Этот принцип постулировал, что интенсивности линий в спектре пропорциональны коэффициентам Фурье «соответствующих» членов разложения электрического момента атома, рассчитанного по классической механике. Эпштейн и Эренфест перенесли принцип соответствия на задачу о дифракции. Они говорят: будем считать, что интенсивность дифракционного спектра m-го порядка определяется, как и в классической волновой теории Релея, квадратом m-го коэффициента Фурье разложения «пропускающей способности» решетки (т. е. поля на ней).

Таким образом, дифракция объяснялась с точки зрения теории световых квантов. Так же обстоит дело и с эффектом Допплера, который объясняется в рамках этих представлений на основании применения законов сохранения [1]. Следовательно, оба явления, достаточные, как показывает рассуждение Бора, для того чтобы в опыте Эйнштейна наблюдалось уширение линии, находят свое объяснение и в теории световых квантов. Таким образом, опыт, предложенный Эйнштейном, не является experimentum crucis для волновой и корпускулярной теории света.

Однако меня не удовлетворяет объяснение Бора, и я считаю, что это не очень глубокое освещение вопроса. На вопрос о том,

[1] *E. Schrödinger.* Phys. Zs., 23, 301, 1922; *A. Sommerfeld.* Atombau und Spektrallinien (изд. 1931 г.), гл. 1, § 7.

что же делается с монохроматическим светом, если его прерывают, теория световых квантов все же не давала ясного ответа. Этот вопрос, мне кажется, нужно было бы рассматривать в связи с квантовой теорией спектральных аппаратов.

ДЕСЯТАЯ ЛЕКЦИЯ

(13.V 1933 г.)

Модуляция частоты. Некоторые задачи из области интерференции света. Задача о распространении радиоволн по поверхности Земли. Оптический вентиль

Рассмотрим вопрос о *модуляции частоты*, который в последнее время приобрел интерес не только для беспроволочной телеграфии, но также и для оптики.

Пусть мы имеем маятник, длина которого (или сила тяжести) периодически и достаточно медленно изменяется с течением времени. Аналогичную систему можно осуществить и для электрических колебаний, если у электрического контура периодически изменять самоиндукцию или емкость. Период маятника, зависящий от длины, будет в этом случае изменяться во времени и, следовательно, колебания не будут гармоническими, так как гармонические колебания — это колебания по закону $a \sin (nt + \gamma)$, где n — постоянная. Если же n, как в нашем случае, является функцией времени, то это уже не гармонические колебания.

Частота колебаний в нашем случае будет $\sqrt{\frac{g}{l}} = f(t)$, она — функция времени. Часто колебания с такой частотой записывают в виде

$$y = a \sin (f(t) \cdot t + \gamma),$$

что, конечно, неправильно. Как же правильно написать выражение для колебаний с переменной частотой? Чтобы решить это, обратимся к школьному изложению вопроса о гармонических колебаниях, как о движении проекции точки, движущейся равномерно по кругу. Координата y выразится, как $y = a \sin \varphi$, где φ — полярный угол точки, a — радиус круга. Число оборотов в единицу времени равно $\dot{\varphi}/2\pi$, а циклическая частота n равна $\dot{\varphi}$. Если $\dot{\varphi}$ постоянна, то $\varphi = nt + \text{const}$ и мы имеем обычное гармоническое колебание.

Если же $\varphi = f(t)$ — функция времени, то

$$\varphi = \int_0^t f(t)\,dt + \text{const}$$

и

$$y = a\sin\left(\int_0^t f(t)\,dt + \text{const}\right).$$

Это и есть выражение для колебания с переменной частотой $n = f(t)$. Мы видим из этого рассуждения, что изображать колебание с этой частотой как $a\sin(f(t)\cdot t + \text{const})$ — это значит делать такую же ошибку, как если для неравномерного движения написать $s = v(t)\cdot t$, а не $s = \int v(t)\,dt$.

Если, например, мы рассматриваем явление Зеемана (или явление Штарка) в переменном магнитном (или переменном электрическом) поле, то частота излучаемого при этом света будет

$$n = n_0 + b\cos\omega t. \quad (59)$$

Это значит, что колебания в этом свете будут

$$y = a\sin\left(n_0 t + \frac{b}{\omega}\sin\omega t + \text{const}\right). \quad (60)$$

Что мы будем наблюдать, если такие колебания действуют на спектральный аппарат, зависит от свойств этого аппарата [1].

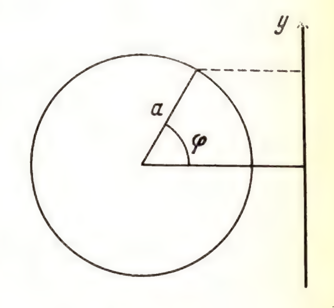


Рис. 23

Вопрос о применении частотной модуляции в радиотехнике одно время стоял довольно остро, так как думали, что частотная модуляция выгоднее, чем амплитудная. При амплитудной модуляции приемник должен принимать довольно широкую область частот. Думали, что, применяя частотную модуляцию, можно сузить необходимую для приема область частот и таким образом «сэкономить место в эфире», т. е. увеличить число работающих без взаимных помех станций в данном частотном интервале. Конечно, это было недоразумением.

Сторонники этого взгляда рассуждали так. Пусть приемник имеет острую резонансную кривую. Тогда очень маленькие изменения частоты передающей станции вызовут большие изменения амплитуды колебаний в приемнике. Но это неверно, так как резонансная кривая характеризует колебательную систему для уста-

[1] См. *М. Дивильковский.* ЖЭТФ, 7, 650, 1937.

новившегося режима. Это часто забывают и упускают из виду скорость установления этого установившегося режима. Чтобы приведенное рассуждение было верно, нужно вести радиопередачу очень медленно, так как если резонансная кривая острая, если прибор селективен, то время установления стационарного режима велико, и это время тем больше, чем прибор селективнее.

Чтобы правильно решить вопрос о том, что будет при действии частотно-модулированного колебания (60) на резонатор, надо разложить это колебание в ряд, в сумму чисто синусоидальных членов. При этом, как и для амплитудной модуляции с частотой ω, в разложении получаются частоты n, $n \pm \omega$, а сверх того и частоты $n \pm 2\omega$, $n \pm 3\omega$, ... Итак, в указанном отношении (ширина спектра) частотная модуляция никакой выгоды не дает.

В первой лекции я указал ряд других вопросов оптики, которые могут привести к парадоксам. Наиболее простые из них — это вопросы, связанные с отражением, преломлением и интерференцией света.

Пусть на тонкую прозрачную пластинку нормально падает плоская волна. Пусть толщина пластинки подобрана так, что волны, отразившиеся от ее верхней и нижней граней, имеют разность хода в полволны ($\lambda/2$). При этом волны, отраженные от верхней и нижней граней, будут уничтожать друг друга, отражения от пластинки не будет, весь свет пройдет. Предположим теперь, что нижняя поверхность пластинки граничит с идеальным проводником (идеальным зеркалом). В этом случае мы получим явное противоречие с законом сохранения энергии: волна падает, но не отражается и не проходит и не абсорбируется.

Здесь дело в том, что для качественного объяснения явления в отсутствие идеального зеркала достаточно было ограничиться двумя интерферирующими лучами. На самом деле число лучей, отражающихся от нижней и верхней поверхностей пластинки, бесконечно велико. Чтобы получить правильный результат, нужно просуммировать действие всех этих лучей.

При наличии же зеркала необходимо учесть все бесконечное число лучей, отражающихся на обеих поверхностях пластинки, так как в этом случае ряд, сумма которого дает действие всех лучей, сходится медленно и первые его два члена уже не дают хорошего приближения. Результирующая амплитуда отраженного света, получающаяся путем суммирования этого ряда (представляющего собой геометрическую прогрессию), равна в этом случае

$$A = \frac{re^{-i\frac{\delta}{2}} - e^{i\frac{\delta}{2}}}{e^{-i\frac{\delta}{2}} - re^{i\frac{\delta}{2}}},$$

где r — коэффициент отражения от первой (верхней) поверхности, а δ — запаздывание по фазе при однократном прохождении пластинки. Амплитуда падающего света принята за единицу. Легко видеть, что $|A| = 1$, т. е. независимо от толщины пластинки весь свет отражается, так что никакого противоречия нет.

Возьмем теперь находящуюся в воздухе прозрачную пластинку такой толщины, чтобы она ничего не отражала (это всегда можно сделать). За ней вплотную поставим вторую такую же пластинку, за второй третью и т. д. до бесконечности. Мы получим полупространство, заполненное стеклом, которое не будет отражать света. Между тем обычно утверждают, что из уравнений Максвелла вытекает наличие отражения на границе двух сред.

Здесь дело в том, что последнее утверждение неправильно. Неправильно, что формулы Френеля для коэффициента отражения получаются *только из уравнений* Максвелла. При их выводе делается *дополнительное* предположение, что имеются только *три* волны: падающая, отраженная и преломленная. В нашем же случае с одной пластинкой и с их совокупностью мы заранее предполагаем наличие *четырех* волн: двух в первой среде и двух во второй. Из различных предположений получаются и разные результаты.

Какое мы имеем основание предполагать при обычном выводе наличие только трех, а не четырех волн? Если мы имеем пластинку из абсолютно непоглощающего материала, абсолютно точно сделанную, гладкую и плоскопараллельную, то физические явления в ней соответствовали бы предположению о наличии четырех волн, как бы толста ни была пластинка. Но поскольку в реальных случаях всегда имеются поглощение и шероховатости, то при большой толщине пластинки реальному случаю отвечает предположение о наличии трех, а не четырех волн.

С только что разобранным вопросом связаны вопросы о роли запаздывающих и опережающих потенциалов в электродинамике. Им в свое время была посвящена полемика между Ритцем и Эйнштейном [1].

Остановимся еще на одном вопросе, связанном с беспроволочной телефонией. Сопоставление распространения радиоволн с вопросами оптики совершенно естественно, ибо речь идет о распространении электромагнитных волн, отличающихся от оптических только своей длиной. Но как раз эта разница и может повести к ошибкам, если формулы из одной области некритически переносить в другую.

Задача о распространении радиоволн вдоль земной поверхности, рассматриваемой как плоская, была поставлена 30 лет

---

[1] *W. Ritz.* Phys. Zs., 9, 903, 1908; *A. Einstein.* Phys. Zs., 10, 224, 1909.

тому назад. Эту довольно сложную задачу решил Зоммерфельд. Почему же понадобилось решение новой сложной теоретической задачи для случая, когда происходит отражение и преломление электромагнитных волн на плоской границе двух сред? Ведь, казалось бы, можно было воспользоваться формулами Френеля, известными из оптики.

Разница здесь в относительных размерах. В то время в радиотелеграфии работали с волнами длиннее 100 *м* и источник излучения находился над землей на расстоянии, малом или сравнимом с длиной волны. В оптике же это не так: формулы Френеля выведены для падения плоской волны и применимы поэтому только в том случае, если расстояние от источника излучения до границы раздела очень велико по сравнению с длиной волны. Если отказаться от этого ограничения, то и получаются формулы Зоммерфельда. Если сделать опыт со световыми волнами в таких условиях, что расстояние источника от границы раздела невелико по сравнению с длиной волны, то получаются результаты, согласные с решением Зоммерфельда [1].

Ряд вопросов возникал в связи с термодинамикой излучения. Релей обратил внимание на приспособление, которое можно назвать световым вентилем [2]. Чтобы понять его устройство, рассмотрим сначала аналогичное устройство, но не обладающее вентильным действием. Возьмем два николя, поставленные на пути луча света так, что их главные плоскости повернуты под углом 45° друг к другу. Между ними поставим слой вещества, обладающего естественным вращением плоскости поляризации, например

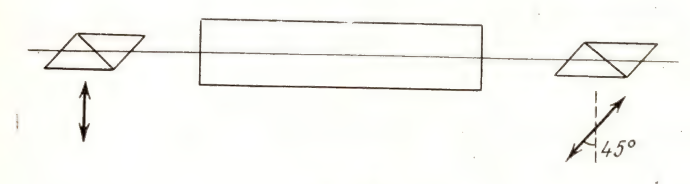


Рис. 24

слой раствора сахара, и возьмем его такой толщины, чтобы поворот плоскости поляризации света, прошедшего сквозь него, составил 45°. Через такое приспособление свет будет проходить так же, как если бы плоскости николей были параллельны друг другу, а слой не обладал бы вращательной способностью. Это будет получаться независимо от направления луча как при про-

---

[1] См. т. I, статья 20.

[2] *Релей.* Sci. Pap., т. IV, стр. 555; Nature, LXIV, 577, 1901.

пускании света в одну сторону, так и при пропускании в другую, так как знак угла поворота плоскости поляризации естественно вращающих тел меняется при изменении направления луча; вращение определяется правилом правого (или левого) винта.

Иначе будет обстоять дело, если вращение плоскости поляризации в слое между николями вызывается не наличием естественной вращательной способности в его веществе, а статическим *магнитным* полем, направленным по лучу, скажем, от первого николя ко второму (явление Фарадея). В этом случае направление поворота плоскости поляризации определяется не направлением луча, а направлением магнитного поля. Поэтому если величина магнитного поля подобрана так, что при распространении света от николя *I* к николю *II* плоскость поляризации повернется на 45° и свет пройдет через николь *II*, то при луче обратного направления, при распространении от николя *II* к николю *I*, плоскость поляризации повернется на 45° в ту же сторону, что и в первом случае, и свет не пройдет через николь *I*.

Такое устройство будет пропускать свет только в одном направлении, оно представляет собой *световой вентиль*.

В. Вин думал, что, используя этот световой вентиль, можно прийти к противоречию со вторым началом термодинамики [1]. Он рассуждал так. Пусть в адиабатической оболочке помещены два одинаково нагретых тела и между ними — световой вентиль.

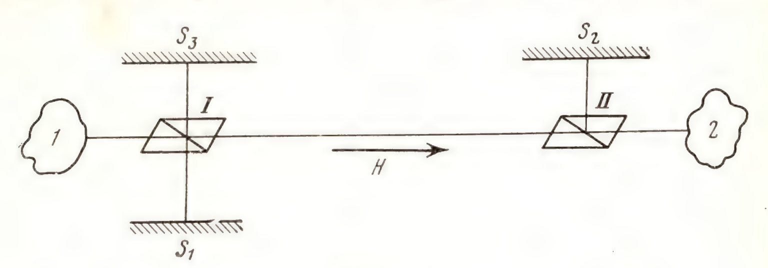


Рис. 25

Свет, излученный телом *1*, падает на николь *I*. Половина падающей энергии J, т. е. $\frac{1}{2}J$, проходит через николь, а другая половина, испытав полное внутреннее отражение в николе, возвращается зеркалом S_1 к телу *1* (применяя зеркало, мы можем не учитывать нагревания николя). Другая половина энергии $\frac{1}{2}J$, прошедшая николь, пройдет затем полностью через николь *II*

[1] *W. Wien.* Rapp. Congr. Int. de Phys. Paris, 1900, т. II, стр. 29.

и поглотится телом *2*. Свет же, посланный телом *2*, не пройдет через световой вентиль. Половина излученной этим телом энергии J, равная $\frac{1}{2}J$, испытав в николе II полное внутреннее отражение и отразившись от зеркала S_2, вернется к телу *2*. Другая половина энергии пройдет николь II, пройдет вращающую среду, испытав при этом поворот плоскости поляризации на 45°, и поэтому не сможет пройти через николь I, а претерпит в нем полное внутреннее отражение. Добавочным зеркалом S_3 она будет отражена обратно и после полного внутреннего отражения в николе I вернется к телу *2*.

Таким образом, тело *1* получит энергию $\frac{1}{2}J$, а тело *2* — энергию $\frac{3}{2}J$ и будет нагреваться в противоречии со вторым началом термодинамики.

Конечно, сам В. Вин не думал, что здесь может нарушаться второе начало термодинамики, но он считал, что компенсация нагревания тела *2* будет получаться за счет каких-то неизвестных процессов, связанных с нагреванием вещества, в котором происходит магнитное вращение плоскости поляризации, так что уменьшения энтропии в противоречии со вторым началом не будет. Однако Релей указал на элементарную ошибку в рассуждении В. Вина[1]. Свет от второго тела, возвращающийся к нему обратно после отражения от николя I, испытает при прохождении вращающегося слоя еще один поворот на 45° и поэтому не пройдет через николь II, а испытав в нем полное внутреннее отражение, вернется (если поставить еще одно зеркало напротив S_2) к телу *1*. Таким образом, второе начало термодинамики не нарушается и нет надобности искать какие-то неизвестные явления, вызывающие компенсацию[2].

[1] *Релей*. Sci. Pap., т. IV, стр. 555, 1901.

[2] [Парадокс В. Вина, таким образом, кажущийся. Все же случай тел с магнитным вращением плоскости поляризации — в известном смысле особый случай в термодинамике излучения. Дело в том, что доказательство законов Кирхгофа для излучения, исключающих возможность противоречий со вторым началом, основано на электродинамической теореме взаимности. Обычное доказательство основано, как известно, на теореме взаимности геометрической оптики (см., например: *Г. Лоренц*. Лекции по теории излучения, гл. I, § 9). Можно, однако, провести вывод законов Кирхгофа, не пользуясь геометрической оптикой, а применяя теорему взаимности в ее общей формулировке. Между тем для тел с магнитным вращением плоскости поляризации теорема взаимности не верна, и поэтому основанное на ней доказательство законов Кирхгофа в этом случае неприменимо.

Законы излучения, заменяющие для гиротропных тел обычные законы Кирхгофа, удалось получить лишь в последнее время, опираясь на так называемую флуктуационно-диссипационную теорему (см. *М. Л. Левин* и *С. М. Рытов*. Теория равновесных тепловых флуктуаций в электродинамике, гл. V. М., Изд-во «Наука», 1967).]

ЛЕКЦИИ ПО ФИЗИЧЕСКИМ ОСНОВАМ ТЕОРИИ ОТНОСИТЕЛЬНОСТИ

(1933—1934 гг.)

ПЕРВАЯ ЛЕКЦИЯ

(7.X 1933 г.)[1]

Введение. Исторический обзор: гипотеза Рёмера, измерения Физо и Фуко, аберрация, Юнг, Френель, опыты Араго и теория Френеля оптических явлений в движущихся телах, частичное увлечение эфира и следствия из него опыт Физо, эффект Допплера, трудности механических теорий эфира

Теория относительности — физическая теория. Эта теория была создана для охвата большого класса физических явлений, которые — надо сказать определенно — долго и упорно не укладывались в рамки существовавших воззрений. Основными здесь были электромагнитные явления в движущихся телах. Настойчивость, с которой занимались этими явлениями, показывает, что в них действительно крылись большие трудности. Эти трудности были связаны с вопросом о взаимодействии между материей и тем, что принято было называть эфиром.

---

[1] [При обработке этого курса лекций основными материалами послужили следующие: 1) записки С. М. Рытова (лекции 1—6), Г. С. Горелика (лекции 1—7), М. А. Дивильковского (лекции 2—6), М. А. Леонтовича (все лекции), З. Г. Либина (лекции 4—6); 2) стенограммы лекций 7—14; 3) конспекты и черновики самого Л. И. Мандельштама, написанные им в процессе подготовки лекций. Л. И. Мандельштам предполагал продолжить курс теории относительности в 1935/36 учебном году и действительно прочел четыре лекции. Как он сам сказал в первой из них, в этом курсе он хотел «коротко повторить историческую часть и релятивистскую кинематику, а затем показать, как строятся релятивистские механика и электродинамика, как они объясняют фактический опытный материал, и если останется время, то несколько слов посвятить общей теории относительности». Однако в четырех прочитанных лекциях (23 октября, 4, 10 и 22 ноября 1935 г.) Л. И. Мандельштам успел только повторить содержание лекций 1—9 курса 1934 г. Записки этих четырех лекций были использованы при обработке помещаемого здесь курса, но сами они, будучи конспективным повторением прочитанного ранее, здесь не помещены. Материал обработан С. М. Рытовым.]

Принцип относительности [1] позволил преодолеть эти трудности и непротиворечиво объяснить явления в движущихся телах. Мы увидим, что релятивистская точка зрения действительно позволяет дать непротиворечивую, стройную и удивительно убедительную теорию этих [механических и электромагнитных] явлений (правда, постольку, поскольку дело идет о макроскопических явлениях), в которую, естественно, укладывается огромный материал, не поддававшийся прежним теориям, и которая была и остается плодотворной основой для всего дальнейшего развития физики вообще. В этом главная и — с точки зрения физики — единственная заслуга принципа относительности. Но это удалось сделать ценой существенного и принципиального изменения, или, вернее, *уточнения* (что часто забывают), некоторых основных понятий, с которыми оперировала физика, и в первую очередь наших воззрений на пространственные и временные измерения. Вот почему принцип относительности перерос те рамки, которые были ему предначертаны непосредственными физическими заданиями. Вот почему им была захвачена и механика, а в конечном счете вся физика. Вот чем объясняется тот, как вы знаете, огромный интерес, который вызвала теория относительности в среде не только физиков.

По существу это здоровый интерес, но формы его удовлетворения нередко уродливы. В литературе, особенно в популярной, все подчас чрезвычайно вульгарировано или заменено непонятным философствованием о непонятных вещах. Для того чтобы действительно понять теорию относительности, нужно, во-первых, знать физику, знать факты, знать те явления, для объяснения которых теория была создана. Во-вторых, речь идет о теории, но сегодня язык физической теории — математика. Нужно владеть тем математическим аппаратом, который адекватен данным проблемам. В настоящее время вряд ли возможно действительно понять и усвоить какую-нибудь теоретическую физическую дисциплину без математики просто потому, что это слишком трудно (это обстоятельство имеет, по моему мнению, глубокое основание). Во всяком случае, в вопросе о принципе относительности я не вижу возможности более или менее серьезной «безматематической» трактовки. Мне кажется, что популярные изложения хотя и дают, может быть, некоторое представление о нем, но главная (физическая) сторона из них ускользает.

И третье. Чтобы усвоить принцип относительности, нужно не бояться думать. Нужно желание передумать «привычные» вещи.

---

[1] [Л. И. Мандельштам сплошь и рядом пользовался термином «принцип относительности» в смысле, эквивалентном «теории относительности», принцип же относительности в более узком смысле он часто называл *постулатом* относительности.]

Усвоение принципа относительности может быть значительно облегчено, если по-настоящему вдуматься в (сравнительно тонкие) основные понятия физики до принципа относительности.

Неправильно полагать, что принцип относительности перевернул наши понятия о времени и пространстве в том смысле, что на место старых ясных и четких понятий он поставил такие же новые. Это не так. Одна из больших заслуг принципа относительности в том и состоит, что он показал, что основные понятия, которыми оперировали раньше — во всяком случае, в известной своей части,— вовсе не были определены, что многие высказывания не имели вообще никакого смысла и что это главным образом и было причиной тех недоразумений, с которыми сталкивались, когда старались подвести теоретическое обоснование под те или иные физические явления. Если мы в этом убедимся, то для нас станет очевидным, что по существу парадоксальность принципа относительности заключается, так сказать, не в его положительных высказываниях, а в том обстоятельстве, что нам не были ясны самые, как казалось, простые и элементарные понятия. Мы увидим далее, что *какое-нибудь* уточнение этих понятий *неизбежно*. И когда мы будем подготовлены к необходимости *какого-нибудь* ответа, то тот ответ, который дает принцип относительности, уже вряд ли нам покажется парадоксальным.

Теория относительности распадается на две части — специальную и общую теорию. Более подробно об этом мы будем говорить потом, когда войдем уже в круг идей принципа относительности. Пока же я ограничусь следующим замечанием.

При исследовании всевозможных движений сразу же выделяется особый, специальный, наиболее простой случай движения — равномерное поступательное движение. Мы хотим изучать законы физических явлений именно в таких системах — движущихся равномерно и прямолинейно. По отношению к чему? По отношению к некоторой системе, являющейся почему-либо выделенной, особой.

Почему именно эта задача всегда привлекала внимание в первую очередь? Исторически ответ ясен. Мы работаем на Земле, которая вращается вокруг оси и движется вокруг Солнца. Периферические скорости суточного вращения гораздо меньше скорости поступательного движения по орбите. В первом приближении движение Земли можно считать именно прямолинейным и равномерным. И принцип относительности создавался, как вы знаете, двумя этапами.

Первый этап завершается в 1905 г., когда Эйнштейн опубликовал свою работу «Zur Elektrodynamik bewegter Körper» [1]. Вы

[1] [Ann. Physik, 17, 891, 1905. Русский перевод в сборнике «Принцип относительности». ОНТИ, 1935.]

видите, какая здесь ставилась, по сути дела, еще скромная задача и как в самом названии работы отразился непосредственно предшествующий круг вопросов — электродинамика в движущейся среде. В этой работе Эйнштейн создал ту часть теории, которая теперь носит название *специального* принципа относительности и которая ограничивается прямолинейным и равномерным движением, т. е. связана с явлениями, происходящими на движущейся Земле. В работе Эйнштейна, несмотря на ее скромное название, заключена вся специальная теория относительности, и в дальнейшем сюда не было добавлено ничего принципиально нового.

Но вся постановка вопроса о принципе относительности вела к тому, чтобы распространить его на любые движения, а тогда неизбежно в круг рассмотрения входит гравитация. Так возникла *общая* теория относительности — второй этап, ознаменованный работой Эйнштейна «Zur allgemeine Relativitätstheorie», вышедшей в 1915 г.[1] В названии этой его работы уже нет речи ни об электродинамике, ни о гравитации, на первом плане — относительность.

Конечно, в наших беседах нет никакой возможности охватить все. Теория относительности пронизала теперь всю физику. Область тех вопросов, в которых она играет роль, необозрима, не говоря уже о содержании самой теории. Мы ограничимся почти без исключений специальной теорией относительности. Разумеется, и здесь есть множество вопросов — и физических, и математических, и философских. Я собираюсь сосредоточить внимание только на первых: на физических основах и на той математической структуре, в которую выливается принцип относительности на базе физических данных. Конечно, и эти вопросы невозможно рассмотреть полностью все. Но это, пожалуй, и не нужно. Я не считаю необходимым останавливаться на отдельных применениях теории и на математических выкладках, конечно, чрезвычайно важных, но, мне кажется, усваиваемых лучше при чтении литературы. Имеются учебники, и учебники отличные, и я вряд ли мог бы прибавить здесь что-нибудь существенное.

И, наконец, последнее замечание. Как правильней будет вести изложение? Вообще говоря, есть два пути: либо излагать теорию в ее современном виде, либо подойти к ней исторически. Мне кажется, дело здесь обстоит так. Знание исторического развития какой-нибудь основной теории всегда интересно и поучительно, но не всегда необходимо. Например, волновую оптику можно излагать без связи с корпускулярной оптикой Ньютона. Но в вопросе о принципе относительности положение, по моему

---

[1] [Berl. Ber., 778, 1915].

мнению, несколько иное, и это по целому ряду причин. Надо исподволь прийти к парадоксальным выводам теории относительности, надо осознать неизбежность этих выводов, надо знать, как пытались обойти трудности крупнейшие ученые и как это действительно не удавалось.

И затем — слишком жива еще связь принципа относительности с именем Лоренца и с предшествующими взглядами, в них уходят его корни, и хотя бы поверхностное знакомство с ними дает, мне кажется, чрезвычайно много для понимания самого принципа относительности.

Позвольте мне поэтому начать с небольшого исторического обзора, причем, конечно, я смогу остановиться только на основных этапах, и то поверхностно.

Основное открытие, явившееся той базой, без которой принцип относительности вообще не существовал бы, было открытие конечной скорости света Олафом Рёмером в 1675 г.

Наблюдая затмения ближайшего спутника Юпитера, время обращения которого в среднем из многих наблюдений было определено Кассини в $1^3/_4$ дня, он обнаружил уменьшение промежутков между затмениями, когда Земля находилась в B, и, наоборот, запаздывание затмений, когда она находилась в A (рис.1). Дело обстояло так, как будто движение спутника возмущено какой-то неизвестной причиной. Рёмер нашел правильное объяснение этому явлению в том, что свет распространяется с конечной скоростью. Он получил значение $3{,}1 \cdot 10^{10}$ *см/сек.*

Факты были настолько убедительны, что Кассини сначала принял это объяснение, невзирая на господствующие взгляды школы Декарта, согласно которым принципиально свет распространяется мгновенно. Однако наблюдения над другими спутниками Юпитера не дали такого же хорошего согласия с теорией Рёмера. Кассини от нее отказался. Сам Рёмер не дожил до подтверждения своей идеи.

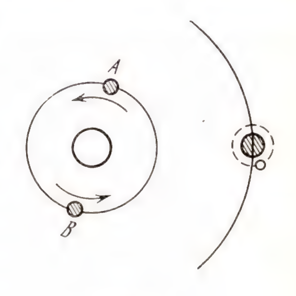


Рис. 1

Особенно настоятельная потребность в определении значения скорости света c возникла в XIX в. В 1849 г. Физо получил $3{,}14 \cdot 10^{10}$ *см/сек*; Фуко измерил c методом вращающегося зеркала (1862 г.); методом Фуко c измерял Корню (1874 г.). Измерение методом Фуко было с очень большой точностью повторено Майкельсоном в 1902 и 1926 гг. Он определял скорость света не только в воздухе, но также в воде и сероуглероде. Для сероуглерода результат оказался отличным от значения, вычисленного из пока-

зателя преломления. Этому было дано следующее объяснение. Скорость света вовсе не такое простое понятие, как казалось вначале. Мы говорим обычно о *фазовой* скорости, но при всех прямых измерениях определяется не фазовая, а *групповая* скорость. Только в вакууме сложное понятие скорости света становится простым, фазовая и групповая скорости совпадают.

В 1926 г. Майкельсон получил наилучший из всех существующих результатов:

$$c = (299796 \pm 1) \cdot 10^5 \ \textit{см/сек}.$$

Я думаю, что эта точность несколько преувеличена. Современные таблицы дают обычно ± 4, но и при такой точности c является наиболее точно измеренной из всех физических постоянных.

Тот факт, что свет распространяется с определенной конечной скоростью, в теории относительности приобрел совершенно исключительное значение. Мы увидим, что он имеет для учения о времени такое же значение, как факт существования твердых тел для учения о пространстве.

Следующее после Рёмера основное открытие, ставшее вместе с тем родоначальником всех последующих трудностей и вопросов оптики движущихся тел, было открытие Джемсом Брадлеем в 1728 г. аберрации. Он сообщает об этом в своем «Отчете о вновь открытом движении неподвижных звезд» [1], присланном королевскому астроному Галлею. Брадлей нашел, что положение звезды γ Дракона меняется в течение года и что это изменение не является результатом параллакса. Он правильно понял, что это явление обусловливается движением Земли по орбите и конечной скоростью света. (Я должен оговориться, что слова «правильно понял» и вообще вся терминология, которую я пока употребляю, — это язык, на котором говорили до 1905 г. Мы находимся пока целиком в кругу дорелятивистских понятий.) Брадлей стоял на точке зрения ньютоновской корпускулярной оптики, и вновь открытое явление — аберрация света — великолепно укладывалось в эту теорию. Действительно, подобное же явление мы наблюдаем, например, из окна железнодорожного вагона при падении дождевых капель.

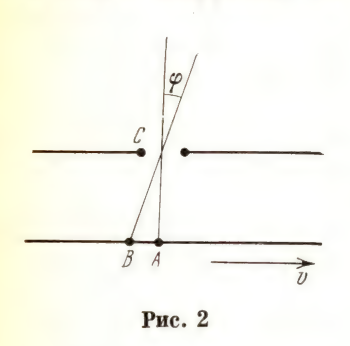


Рис. 2

Если мы смотрим на звезду через диафрагму C (рис. 2), то вследствие движения Земли световая корпускула, летящая от

[1] См. *E. T. Whittaker*. A History of the Theories of Aether and Electricity. Dublin, 1910, стр. 99.

звезды по линии CA, попадет не в точку A, а в точку B, которая станет на место A к моменту падения корпускулы на прямую AB. Для угла аберрации φ — угла между направлением AC на звезду при покоящейся Земле и направлением BC при движущейся Земле — получается, очевидно, $\operatorname{tg} \varphi = \frac{v}{c}$. Отсюда Брадлей мог определить скорость света. Он нашел для времени распространения света от Солнца до Земли значение 8 мин. 12 сек., что совпадало со средним из значений, определенных по методу Рёмера из затмений спутника Юпитера. Конечно, нужно было предположить, что сама Земля не действует заметно на частицы света. Замечу, что обнаружить аберрацию вообще можно лишь благодаря изменению направления движения Земли. Без этого мы не знали бы истинного направления на звезду.

После Брадлея в XVIII в. не было ни наблюдений, ни теоретических исследований по интересующим нас вопросам. Уиттэкер говорит о бесплодии этого периода [1]. Физики XVIII в. почти все были приверженцами ньютоновской корпускулярной теории. Волновой точки зрения, выдвинутой Гюйгенсом в его знаменитом «Traité de la lumière» еще в 1690 г., придерживались лишь немногие, в частности Франклин и Эйлер; последний подчеркивал сходство света со звуком [2]. Положение начало меняться лишь в самом конце XVIII в. с появлением Юнга. Ведущие ученые, возглавлявшие в это время французскую Академию наук, Лаплас, Пуассон, Био по-прежнему стояли на корпускулярной точке зрения, и последняя имела даже некоторые достижения. Так, например, Лаплас сумел довольно хорошо объяснить двойное лучепреломление.

В 1799 г. Юнг выступил с критикой господствующей теории. Он говорит, что корпускулярная теория его не удовлетворяет, ибо он не понимает, как она может объяснить установленный с огромной точностью факт независимости скорости света от его интенсивности... Волновая же теория объясняет это совершенно естественно. Трудно сказать, почему именно *это* было непонятно Юнгу, а не что-нибудь другое, но с этого он начинает свою критику.

Юнг — вообще замечательный человек. Оказывается, он еще тогда высказал мысль о существовании электрического эфира и о возможном тождестве его с эфиром световым [3]. «Но,— пишет Юнг,— я до сих пор не смог обнаружить, чтобы преломляющая сила эфира претерпевала какое-либо изменение под действием электричества». Юнг первый дал правильное объяснение интер-

[1] Loc cit., стр. 100.
[2] Loc. cit., стр. 103.
[3] Loc. cit., стр 105.

ференции, он первый истолковал явления поляризации как результат поперечности световых колебаний. Надо иметь в виду, что это был чрезвычайно трудный шаг, ибо в то время теории упругости твердых тел еще не существовало, а в газах и жидкостях возможны лишь продольные волны.

В 1804 г. Юнг дает объяснение аберрации с точки зрения волновой теории. Это общеизвестное элементарное (кустарное) объяснение, но все же действительно объяснение. Оно основано на предположении, что эфир *не увлекается* Землей. Тогда волны, прошедшие через диафрагму (объектив трубы) и распространяющиеся в эфире так же, как если бы Земля покоилась, попадут левее того места AB, куда они пришли бы в случае неподвижной Земли (см. рис. 3). Равным образом в том случае, если эфир полностью увлекается Землей, волны попадут на AB, т. е. никакой аберрации не будет.

Мы видим, что весь вопрос приобретает принципиальный интерес, что здесь сразу же возникает та кардинальная проблема, которая потом оставалась не разрешенной вплоть до принципа относительности — проблема взаимодействия между движущейся Землей и эфиром, а тем самым и проблема взаимодействия между эфиром и материей вообще [1].

Несмотря на мощный толчок, который получила волновая теория со стороны Юнга, корпускулярная теория не сдавалась. Дифракция света оставалась камнем преткновения для волновой теории, и, как говорит Уиттэкер, сторонники корпускул решили, что настало время для окончательной победы их теории [2]. И вот Парижская академия наук предложила «дифракцию» в качестве темы на премию на 1818 г. Однако надежды не оправдались: на конкурс была представлена работа, в которой с волновой точки зрения объяснялись *все* известные интерференционные, дифракционные и поляризационные явления. Автором этой работы был Френель.

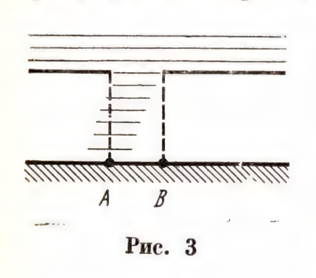


Рис. 3

Тогда и разыгрался знаменитый инцидент между Пуассоном и Френелем. Пуассон, бывший членом конкурсной комиссии, обратил внимание на «нелепый» вывод, вытекающий из теории Френеля: в центре тени круглого диска должно находиться светлое пятно. Араго тут же сделал опыт, и оказалось, что такое пятно

[1] [Л. И. Мандельштам везде пользовался термином «материя» в смысле старого термина «весомая материя», т. е. вещество.]

[2] Loc. cit., стр. 113.

действительно есть. Победа волновой теории была окончательной и ее признание — всеобщим.

В том же году Френель напечатал работу о влиянии движения Земли на оптические явления, в которой он высказался определенным образом и по вопросу о взаимодействии материи и эфира. Еще до этого Мичель обратил внимание на то, что при неувлекаемом эфире коэффициент преломления тела должен быть различен в зависимости от того, движется ли тело навстречу источнику света (скорость световых частиц, падающих на тело, равна при этом $c + w$) или удаляется от него (скорость равна $c - w$). Следовательно, должно меняться и фокусное расстояние линзы, и угол отклонения луча призмой. Боскович указал на другие аналогичные явления.

И вот Араго, который стоял тогда еще на корпускулярной точке зрения, делает следующий опыт. Он смотрит, изменяется ли показатель преломления призмы для света от звезды благодаря движению Земли. Ожидаемое изменение угла отклонения луча составляло $1'$. Араго не получает никакого эффекта и делает заключение, что движение Земли на показатель преломления не влияет. Это заключение правильно, действительно не влияет. Но доказал ли это Араго своими опытами? Он не мог этого доказать: точность его опытов была недостаточна, чтобы обнаружить столь малый эффект, и он, по сути дела,— как это часто бывало в истории физики — *угадал* правильный вывод, интуитивно почувствовал, что зависимости от движения Земли нет.

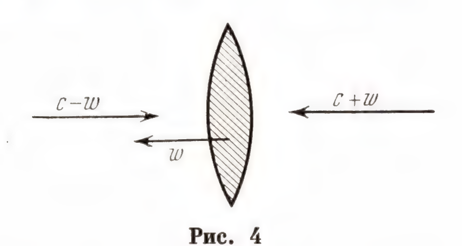


Рис. 4

Так или иначе, этот отрицательный результат считался фактом и требовал объяснения. Сам Араго, насколько можно судить, видел только следующий, правда довольно остроумный, выход из положения, который он излагает в своем письме к Френелю в 1818 г. Предположим, пишет Араго, что свет состоит из частиц, вылетающих из источника со всевозможными скоростями, но наш глаз чувствителен только к частицам со скоростью c. Тогда в случае движущейся призмы (т. е. призмы, движущейся вместе с нами со скоростью w к источнику) мы увидим те час-

тицы, скорость которых относительно источника равна $c - w$, а скорость относительно призмы равна $(c - w) + w = c$, и, таким образом, никакого влияния движения призмы не обнаружим. Однако Араго не был удовлетворен этим объяснением, и в конце письма он обращается к Френелю с просьбой подумать, нельзя ли объяснить отсутствие эффекта с волновой точки зрения.

В ответном письме Френель дает свое объяснение. Гипотезу Араго он считает действительно неудовлетворительной, ибо трудно допустить столь резкую избирательность нашего глаза, что малейшее отклонение скорости от c уже делает свет невоспринимаемым. И он дает свою теорию вопроса, теорию волновую, основанную на предположении о том, что все пространство заполнено *неподвижным* эфиром. Кстати говоря, что понимали под словом «неподвижный»? Говоря об *абсолютном* движении Земли, физики имели в виду ее движение по отношению к «неподвижным» звездам. Таким образом, «неподвижный эфир» означало неподвижный в системе неподвижных звезд.

Теория Френеля оказала громадное влияние на все последующее. В его письме содержится предсказание «частичного увлечения» эфира — и посейчас одного из наиболее важных явлений в движущихся телах. Всякая теория, в том числе и принцип относительности, должна объяснить это явление.

Френель говорит так: действительно, скорость света в стекле иная, чем в пустоте, но само стекло здесь, так сказать, *ни при чем*. Дело просто в том, что стекло, как и другие тела, *сгущает* заключенный в нем эфир. Плотность эфира в стекле (ρ_1) больше, чем в пустоте (ρ), упругость же эфира неизменна (модуль сдвига N). Так как скорость распространения поперечных колебаний равна $\sqrt{\frac{N}{\rho}}$, то коэффициент преломления в случае неподвижного стекла есть

$$n = \frac{c}{c_1} = \sqrt{\frac{N}{\rho}} : \sqrt{\frac{N}{\rho_1}} = \sqrt{\frac{\rho_1}{\rho}}. \tag{1}$$

Пусть теперь стекло движется по отношению к эфиру со скоростью w, т. е. эфир *вне стекла* обтекает его со скоростью $-w$. Но скорость эфира по отношению к стеклу *внутри стекла* будет иной, так как плотность эфира там другая, а мы требуем *сплошности*, неразрывности эфира на поверхности S (рис. 5). Другими словами, количество эфира, притекающее ежесекундно извне на единицу поверхности S, т. е. ρw, должно равняться количеству эфира, оттекающего от этой единицы поверхности за то же время внутрь, т. е. $\rho_1 w_1$:

$$\rho w = \rho_1 w_1.$$

Отсюда и из (1) получаем

$$w_1 = w\frac{\rho}{\rho_1} = \frac{w}{n^2}.$$

Таким образом, внутри, где эфир уплотнен, его скорость по отношению к стеклу есть $w_1 < w$, а по отношению к внешнему эфиру она равна, следовательно,

$$w - w_1 = w\left(1 - \frac{1}{n^2}\right). \tag{2}$$

Выражение $1 - \frac{1}{n^2}$ и есть знаменитый френелевский *коэффициент увлечения*. Он означает, что эфир *увлекается* стеклом ($w_1 \neq w$), но не полностью ($w_1 \neq 0$), а *частично*. Конечно, мы модернизовали изложение Френеля. Стокс писал, что у Френеля нет соображений о сплошности эфира. Лармор считал, что вообще у него не все понятно. Дело в том, что сам Френель пишет о том, что стеклом увлекается (полностью) лишь *избыток плотности* эфира в стекле... Именно эта трактовка считалась туманной и непонятной. Разумеется, взгляды Френеля далеки от современных, но просто поразительно, как по-своему он сумел предугадать дальнейшее развитие вопроса, а именно то его освещение, которое он получил потом в электронной теории.

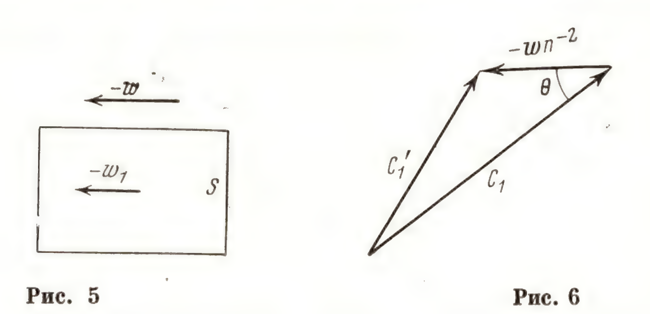


Рис. 5 Рис. 6

Имея (2), нетрудно ответить и на следующий вопрос: какова скорость света в стекле по отношению к наблюдателю, движущемуся вместе со стеклом? Вопрос этот чисто кинематический и решается сразу. По отношению к эфиру внутри стекла свет распространяется со скоростью c_1, но сам эфир движется относительно стекла со скоростью $-\mathrm{w}n^{-2}$. Скорость c_1' равна просто геометрической сумме $\mathbf{c}_1 - \mathrm{w}n^{-2}$ (рис. 6), откуда с точностью до членов второго порядка относительно $\frac{w}{c_1}$

$$c_1' = c_1 - wn^{-2}\cos\theta.$$

Конечно, можно найти много возражений против вывода Френеля. Можно спросить, например, почему он берет различной в стекле и в пустоте *плотность* эфира, а не его упругость, что совершенно изменило бы дело, и т. д. Но так или иначе результат получается у него верный. Действительно, говорит Френель, угол преломления зависит от w из-за изменения скорости падающей волны, но этот эффект *в первом порядке компенсируется эффектом частичного увлечения.* Таким образом, ответ Френеля совершенно иной, чем у Араго.

Резюмируем. Эфир существует и неподвижен. Благодаря этому преломление света в движущемся теле должно отличаться от преломления в неподвижном. Опыт говорит, что этого нет. Френель возражает: есть, но эффект этот очень мал, так как в первом порядке он компенсирован частичным увлечением. Это утверждение мы теперь и докажем. Заметим, что при *полном* увлечении эфира оптические явления, конечно, протекали бы в точности одинаково как для движущегося, так и для неподвижного наблюдателя. Но тогда не было бы аберрации.

Мы докажем даже гораздо больше, а именно, что движение относительно эфира не оказывает влияния на *любые* оптические явления с точностью *до* второго порядка относительно $\frac{w}{c}$, т. е. с точностью, которая только и была доступна во времена Френеля и еще долгое время после него. Я приведу доказательство Френеля в той модернизированной форме, которую дал ему Лоренц.

Пусть луч света распространяется в некоторой среде, показатель преломления которой n есть функция точки, прерывная или непрерывная — безразлично. Если эта среда покоится, то время, в течение которого свет идет из A в B, выразится интегралом

$$\tau = \int_A^B \frac{ds}{c_1},$$

где $c_1 = c_1(x, y, z)$. Если же среда движется со скоростью w вдоль оси x, то вместо c_1 нужно взять c_1', т. е. с точностью до $\frac{w^2}{c^2}$ время прохождения из A в B будет

$$\tau' = \int_A^B \frac{ds}{c_1'} = \int_A^B \frac{ds}{c_1 - wn^{-2}\cos\theta} \approx \int_A^B \frac{ds}{c_1} + \frac{w}{c^2}\int_A^B \cos\theta\, ds = \tau + \frac{w}{c^2}(x_B - x_A).$$

Таким образом, *время прохождения меняется уже в первом порядке,* но это изменение *зависит лишь от положения начальной и конечной точек и не зависит от формы пути.*

Как известно, вся геометрическая оптика может быть основана на принципе Ферма: из всех возможных путей луч пойдет по тому, который дает экстремальное значение τ (в нашем случае линзы или призмы действительный путь даст $\tau = \min$). Отыскивая действительный путь луча из A в B, мы в случае неподвижной среды будем требовать $\tau = \min$, а в случае движущейся среды — $\tau' = \min$. Но второй член в τ' не зависит от формы пути и, следовательно, никак не скажется на виде экстремали. А это значит, что *ход лучей, вся геометрическая оптика, в первом порядке будут для покоящейся и для движущейся среды одни и те же.* Это верно лишь в первом порядке, но ведь только на такую точность и могли претендовать все тогдашние опыты.

Однако это еще не все. Легко видеть, что *интерференционные и дифракционные явления в первом порядке тоже не зависят от движения среды.* Действительно, во всех интерференционных явлениях имеются два пути света и явление зависит лишь от *разности* времен пробега лучей. Но добавочный член в τ' не зависит от формы пути, и потому из разностей времен он выпадает. Тем самым все оптические опыты, в которых путь света замкнут (опыты с одними часами), могут обнаружить движение относительно эфира только во втором порядке относительно $\frac{w}{c}$. Однако на времени пробега в *одном* направлении движение по отношению

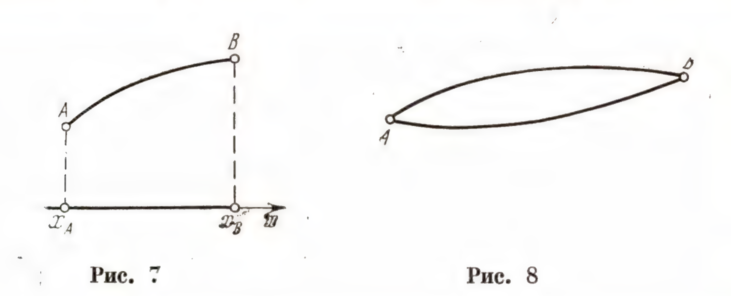


Рис. 7 Рис. 8

к эфиру должно сказаться уже в первом порядке. Таким образом, опыт, в котором измерялось бы самое время пробега (опыт с двумя часами), дал бы отличие уже в порядке $\frac{w}{c}$.

Один такой опыт мы знаем. Это опыт Рёмера, и еще Максвелл предлагал решить с его помощью вопрос об увлечении или неувлечении эфира Землей. До теории относительности считали, что опыт неосуществим лишь из-за недостаточной точности измерений, и не замечали того, что он попросту лишен смысла. Не вдаваясь сейчас в детальное обсуждение, замечу лишь следующее

Опыт аналогичен известному опыту Штурма и Колладона, которые измеряли скорость звука в воде на Женевском озере. В одном пункте производился звук и одновременно давалась световая вспышка, а в другом пункте измерялось время между вспышкой и приходом звука. Но что будет служить такой вспышкой, если измеряется скорость света? Именно здесь кроется вся трудность, в этом — гвоздь вопроса, хотя дорелятивистская физика не видела здесь ничего принципиального.

В 1851 г. Физо подтвердил френелевскую формулу коэффициента увлечения опытами с распространением света в движущейся воде. Майкельсон и Морлей (1878 г.) также нашли полное согласие с формулой Френеля. Не так давно коэффициент увлечения был очень точно измерен Зееманом. Его устройство несколько отличалось от майкельсоновского, и соответственно была переработана теория. Результат и здесь дал полное подтверждение формулы Френеля. Замечу, что в теории относительности френелевский «коэффициент увлечения» является просто следствием релятивистского закона сложения скоростей.

Нужно указать еще на одно из величайших открытий — эффект Допплера. В 1842 г. Христиан Допплер теоретически показал, что цвет света должен зависеть от движения наблюдателя и источника. Наблюдаемый период колебаний T' связан с истинным периодом T соотношением

$$\frac{T'}{T} = \frac{c - w_{\text{ист}}}{c - w_{\text{набл}}}, \tag{3}$$

где $w_{\text{ист}}$ и $w_{\text{набл}}$ — скорости вдоль по лучу. Если наблюдатель и источник движутся вместе, то $T' = T$, но во втором порядке есть разница между случаем, когда наблюдатель приближается к неподвижному источнику $[T'/T = c/(c + w)]$, и случаем, когда источник приближается к неподвижному наблюдателю $[T'/T = (c - w)/c]$ (рис. 9).

Эффект Допплера является прямым следствием конечной скорости распространения света и хорошо укладывается в теорию неувлекаемого эфира. В этом предположении формула (3) выводится элементарно. Физо первый (в 1848 г.) указал на громадное значение эффекта Допплера для астрономии, для определения лучевых скоростей звезд независимо от величины расстояния до них [1].

Позвольте мне резюмировать. С одной стороны, имеется аберрация, с другой — независимость оптических явлений от дви-

[1] [Исследование оптического эффекта Допплера в лабораторных условиях (с помощью наземного источника света) впервые было проведено в 1894—1900 гг. А. А. Белопольским.]

жения Земли. Правда, независимость лишь в первом порядке, но еще не было опытов, которые могли бы пойти дальше. Ведь для Земли $\frac{w^2}{c^2} \sim 10^{-8}$. Френель показал, как совместить в первом порядке результаты наземных опытов с аберрацией, и в течение некоторого времени царило благополучие. Но потом, не столько даже из-за новых фактов, сколько в результате углубленного исследования, начались неприятности. Эфир, когда за него взялись всерьез и начали вникать во все детали, стал создавать одну трудность за другой.

Для возможности поперечных колебаний надо принять, что эфир — твердое тело. Но как примирить такое представление с тем, что все планеты свободно проходят сквозь него? Согласившись все же с тем, что эфир твердое тело, мы сталкиваемся с рядом трудностей в вопросе об отражении света. На границе раздела упругих сред должно соблюдаться *шесть* условий непрерывности — для трех скоростей и трех напряжений. Поперечные же колебания могут удовлетворить только *четырем* условиям. Еще в 1837 г. Стокс показал, что отражение и преломление поперечных волн в упругих средах неизбежно должно сопровождаться возникновением продольных волн. Но в оптике таких волн нет.

Другой вопрос — о плоскости поляризации. Френель, исходя из своей теории эфира, обладающего всюду одинаковой упругостью, считал, что колебания происходят перпендикулярно к плоскости поляризации. Нейман доказывал, что колебания совершаются параллельно плоскости поляризации, так как он принимал постоянной плотность эфира, а различной его упругость.



Рис. 9

Делал он это неспроста. По Френелю получается, что в *анизотропных* телах различна по разным направлениям *плотность* эфира. Но ведь она все-таки скаляр...

Стокс полагал, что ему удалось доказать возможность аберрации в случае увлекаемого эфира в предположении, что он несжимаем и движение его безвихревое. Однако впоследствии Лоренц показал, что это неверно. При наличии потенциала скоростей несжимаемая жидкость неизбежно должна скользить по поверхности движущегося в ней шара. Это скольжение достигает 1,5 w.

Планк попытался спасти аберрацию в увлекаемом эфире, допустив, что он сжимаем. Тогда плотность его у поверхности Земли должна быть в e^{11} раз больше, чем в межзвездном пространстве. Наряду с этим скорость распространения света c и тут и там одинакова. Подобные вещи могут нравиться или не нравиться, но удовлетворить они не могут.

Возникли такие трудности, что эфир вообще утерял что-либо общее с обычными телами. Большое количество попыток дать механическую теорию эфира предпринимал Виллиам Томсон. Так, например, он указал на возможную аналогию со смолой: она течет при медленных движениях (отсутствие сопротивления эфира движению планет), но камертон, сделанный из смолы, звучит, т. е. для быстрых колебаний она ведет себя как твердое тело (распространение в эфире поперечных световых колебаний). Однако с аналогиями много не сделаешь. Количественно все это не получалось. Ведь нужно было, чтобы сопротивление движению планет было не «очень мало», а *совсем отсутствовало*. В этом отношении вопрос напоминает проблему сверхпроводимости, где тоже нельзя построить теорию приближенно (т. е. допустить «очень малое» сопротивление). Таким образом, теория эфира как упругого тела зашла в тупик. Удовлетворительного разрешения проблемы не было.

В 60-х годах весь вопрос получил совершенно другую постановку. В эти годы Максвелл в ряде работ создал свое учение об электрических и магнитных явлениях и дал электромагнитную теорию света. Наступил перелом. Как обстояло дело со всеми этими вопросами в теории Максвелла — об этом в следующий раз.

ВТОРАЯ ЛЕКЦИЯ

(19.XII 1933 г.)

Краткое резюме. Исторический обзор (продолжение): электромагнитная теория Максвелла, проблема электродинамики движущихся тел, опыты Роуланда, Рентгена, Эйхенвальда и Вильсона, теория Герца, ее недостатки, формулировка задачи о явлениях в движущихся телах, вопрос о пересчете от одной системы отсчета к другой в механике

Резюмируем сказанное в прошлый раз. По прошествии продолжительного времени после открытия Рёмером конечности скорости света и открытия Брадлеем явления аберрации интерес привлекли оптические явления в движущихся телах. С точки зрения механической волновой оптики здесь особенно резко выявлялось взаимодействие между материей и эфиром, здесь открывалась возможность изучать свойства эфира, а это всегда привлекало исследователей. Ряд опытов показывал, что эфир, заполняющий все межпланетное пространство, покоится, не увлекается Землей, а отсюда следовало, что движущиеся тела должны испытывать «эфирный ветер». Но большое число опытов со всей доступной в то время точностью свидетельствовало о том, что никакого «эфирного ветра» нет. Это трудно было примирить с существованием аберрации. Выход указал Френель.

Он принял, что тела увлекают эфир совершенно специальным образом, а именно: скорость света в теле, движущемся относительно эфира со скоростью w, составляет по отношению к этому телу

$$c_1' = c_1 - wn^{-2}.$$

Таким образом, «эфирный ветер» в принципе существует, но в тех явлениях, которые обычно наблюдают и в которых измеряется разность времени пробега или свет проходит туда и обратно (т. е. путь света замкнут), в первом порядке относительно $\frac{w}{c}$ получается компенсация влияния движения тела. Это влияние может быть обнаружено лишь во втором порядке, т. е. для Земли — в порядке $\frac{w^2}{c^2} \approx 10^{-8}$, что лежало далеко за пределами достижимой в то время точности. Правда, был один опыт, будто бы противоречащий теории Френеля. Физо в 1860 г. полагал, что им обнаружено изменение поворота плоскости поляризации при прохождении света через стопку стеклянных пластин, изменение, зависящее от движения Земли. Но, как показал в последнее время Штрассер, Физо просто ошибался.

Френель первый подчеркнул, что в опытах Араго с лучом от звезды и во всех других подобных опытах безразлично, делать ли опыт со звездным светом или с земными источниками, так как с точки зрения волновой теории (в отличие от корпускулярной) скорость света не зависит от движения источника. Замечу, что это одно из основных положений, на которые опирается теория относительности. У Френеля оно было гипотезой, высказанной по аналогии с тем, что было известно о скорости звука. Но теперь мы думаем, что это действительно так.

Когда всерьез взялись за толкование эфира как тела, то ввиду поперечности световых колебаний пришли к необходимости сделать его твердым телом. При этом натолкнулись на трудности в вопросе о свободном прохождении планет, в вопросе об отражении, о преломлении, о двойном лучепреломлении и т. д. Пытаясь обойти эти трудности, строили эфиры с очень интересными свойствами, создавали чрезвычайно остроумные теории эфира. Но тем временем произошло событие, одно из самых больших событий в истории физики, которое все изменило и направило интерес в другую сторону. Максвелл провозгласил электромагнитную теорию света.

Максвелл занимался не оптикой, а электродинамикой. Он стремился к тому, чтобы дать полную теорию электромагнитных явлений. Разработанная теория существовала и до него, но он ввел ток смещения.

Сказать так, это, конечно, еще ничего не сказать.

Уравнения, к которым он пришел, не только охватывали все известные электромагнитные явления, но давали больше. Из них вытекало, что всякое электромагнитное возмущение распространяется в виде поперечной волны с конечной скоростью. Таких волн не знали, но теория Максвелла позволяла даже *предвычислить скорость их распространения.*

Уже в прежние теории электромагнетизма входил параметр, который связывал электрические явления с магнитными и который равнялся отношению электромагнитной единицы заряда к электростатической. Мерой заряда может служить сила притяжения пластин конденсатора. Электростатическую единицу заряда определяют тем, что в законе Кулона

$$F = k\frac{e_1 e_2}{r^2}$$

полагают $k = 1$. С другой стороны, между токами тоже существуют силы взаимодействия. Сила взаимодействия двух параллельных элементов ds_1 и ds_2 токов i_1 и i_2 равна

$$F = \beta\frac{i_1 i_2}{r^2}\,ds_1 ds_2.$$

Кроме того, рассматривая изменение заряда обкладок конденсатора при его зарядке, мы можем определить ток как $i = \frac{de}{dt}$. Тем самым константа β связывает механические действия текущего электричества с механическими действиями статического. Сравнивая те и другие силы, можно определить β. Эта величина имеет размерность $\frac{1}{\text{скорость}^2}$, и значение соответствующей «скорости» на основании чисто электромагнитных измерений Вебера и Кольрауша оказалось равным 310 000 *км/сек*. Однако ничего общего с какой-либо действительной скоростью эта «скорость» не имела.

Правда, Кирхгоф связал этот параметр с реальной скоростью. Еще из домаксвелловской теории он нашел, что именно с такой скоростью должно происходить волнообразное распространение тока и заряда вдоль кабеля. Все были поражены тогда совпадением полученной скорости со скоростью света, но связать их между собой не представлялось возможным и совпадение казалось случайным.

Из уравнений Максвелла следовало, что возможны электромагнитные волны и что скорость их распространения как раз равна тому параметру c, который был измерен Вебером и Кольраушем посредством чисто электрического опыта. По-видимому, именно это и заставило Максвелла прийти к выводу, что волны света — электромагнитные волны.

Маленькая деталь, сообщаемая Джинсом: Максвелл опубликовал свою теорию в 64—65 гг.; в письме к Фарадею (1861 г.) он пишет, что из данных Вебера и Кольрауша он вычислил упругость эфира в воздухе и нашел, что скорость электромагнитных волн должна равняться 193 088 *милям* в секунду, в то время как Физо получил непосредственным измерением для скорости света 193 118 *миль* в секунду. Откуда он взял эти числа — неизвестно. В опубликованной теории он приводит другие числа — 310 700 и 314 558 *км/сек* (Физо), дающие гораздо худшее совпадение. Хорошее совпадение, случайно существовавшее между результатами, первоначально известными Максвеллу, несомненно, сыграло чрезвычайно большую роль, ускорив созревание его гипотезы. Обратное случилось с Ньютоном, когда он, располагая неточным значением радиуса Земли, получил для притяжения Луны неправильный результат и на много лет забросил свою теорию всемирного тяготения.

Идеи Максвелла не снимали вопроса о механической теории эфира. Наоборот, проблема расширялась, так как с этих пор уже не было оптических и электрических явлений, не было светового и электромагнитного эфира. Существует один общий эфир, опти-

ка сделалась частью электродинамики, и открылась возможность судить об эфире на основании электромагнитных опытов. Таким образом, вопрос, как сказано, не был снят, а был усложнен, и прошло еще немало времени, прежде чем механические конструкции эфира потеряли интерес. Нужно было объяснить не только волны, но и все остальные электромагнитные явления. Появились работы, в которых эфир рассматривался уже даже не как среда, а как *машина*, строились модели с колесами и т. п. При этом возникали порой комичные трудности: если, например, связать электрическое поле с поступательным движением, то магнитное будет соответствовать вращению. Но тогда при переходе от точки к точке неизбежен разрыв сплошности. Чтобы обойти эту трудность, вводили промежуточные передаточные колеса (рис. 10).

Как замечает Борн, логически все это не могло иметь успеха, так как развитие теории шло в обратном направлении — к тому, чтобы механические свойства тел объяснить электрическим взаимодействием частиц. Но зачем же тогда опять сводить это взаимодействие к механике? В конце концов стали смотреть на эфир так, что он вполне характеризуется своими электромагнитными свойствами и что вредно требовать от него еще и механических свойств.

Не следует думать, что электромагнитная теория света сразу же завоевала себе признание. В вопросе об отражении сам Максвелл натолкнулся на затруднения, которые лишь потом были полностью устранены Герцем. Но главное заключалось в том, что

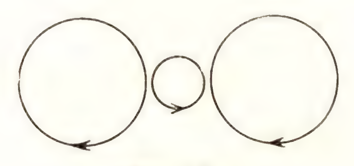


Рис. 10

электромагнитных волн не знали и до опытов Герца теория Максвелла оставалась предположением. (Первоначально Герц допустил ошибку в вычислениях и получил измеренное значение скорости электромагнитных волн на 40% меньше теоретического. Ошибка была указана Пуанкаре, и совпадение получилось прекрасное.)

После опытов Герца положение дела изменилось. Однако В. Томсон еще долгое время не мог примириться с электромагнитной теорией света. В своих лекциях в 1904 г. он говорит, что «так называемая электромагнитная теория света до сих пор нам

ни в чем не помогла»[1]. Надо было быть гением (или наоборот), чтобы не соглашаться с электромагнитной теорией... Томсону это было позволительно. К началу этого столетия электромагнитная теория была уже всеми признана.

Нас интересует то обстоятельство, что в вопросе о явлениях в движущихся телах вышли на сцену электромагнитные явления. Конечно, частные вопросы такого рода возникали и раньше. Еще в 1838 г. Фарадей писал, что если посреди комнаты подвесить заряженный шар и потом заставить его двигаться в каком-либо направлении, то эффект будет тот же, как если бы мы возбудили ток в направлении движения. Однако Фарадей не смог доказать этого на опыте. Лишь в 1876 г. такой опыт был по предложению Гельмгольца сделан Роуландом[2].

Роуланд взял два позолоченных стеклянных диска (рис. 11), между которыми вращался оклеенный золотой фольгой эбонитовый диск. Обкладки на эбоните заряжались, скажем, положительно, а обкладки на стекле заземлялись. Такое устройство — в виде конденсаторов — позволило работать с большими зарядами. Астатическая магнитная стрелка была подвешена над верхней стеклянной крышкой, и при вращении эбонитового диска наблюдалось отклонение этой стрелки. Этот опыт очень точен.

Несколько лет спустя (1888 г.) Рентген сделал другой опыт — с поляризованным диэлектриком[3]. Между разноименно заряженными обкладками вращался диск из незаряженного изолятора (рис. 12). Отклонение стрелок магнетометра показывало, что и в этом случае возникает ток. Этот опыт очень труден и первоначально встретил возражения (Роуланд указывал на возможность

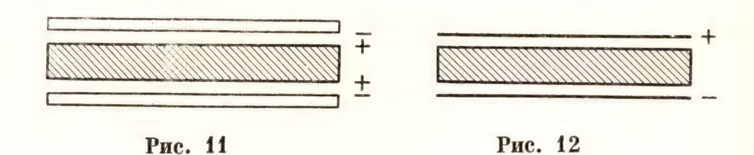


Рис. 11 **Рис. 12**

утечки зарядов), но Рентген все их опроверг. Рентген добивался высокой чувствительности устройства для того, чтобы решить вопрос, увлекается ли эфир полностью или же так, как следует по Френелю. Однако он не смог достичь необходимой точности. Количественные результаты были достигнуты Эйхенвальдом в Москве в 1904 г. и много дали для теории относительности.

---

[1] *Whittaker*, стр. 297.
[2] Loc. cit., стр. 339
[3] Loc. cit., стр. 426.

Вильсоном был сделан еще и такой опыт: он заставлял вращаться полый цилиндр из незаряженного диэлектрика в магнитном поле, направленном по оси (рис. 13). К металлическим обкладкам цилиндра присоединялся электрометр, который показывал, что при вращении цилиндра между обкладками появляется напряжение.

Что же говорят нам эти опыты? Какие выводы можно из них сделать?

Чтобы ответить на это, надо иметь теорию электромагнитных явлений в *движущихся телах*. Уравнения Максвелла не содержат в себе такой теории, ибо *они относятся к неподвижным телам*. Таким образом, возник вопрос о распространении этих уравнений на движущиеся тела. При этом надо было помнить, с одной стороны, об оптических вопросах, а с другой — о том, что вся электротехника основана на движении одних тел по отношению к другим (динамомашины, моторы). Это во-первых. Во-вторых, мы живем на Земле, и это обстоятельство выдвигает вопрос о том, как здесь протекают явления, т. е. как влияет на них равномерное прямолинейное движение.

Я уже здесь хотел бы обратить ваше внимание на следующее. На первый взгляд постановка вопроса кажется ясной. Но если немного глубже в нее вдуматься, то в ней самой обнаруживается известная трудность. Действительно, что значит покоящееся тело и тело движущееся? Относительно чего?

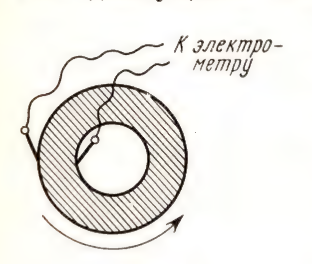


Рис. 13

Мы пока не будем глубже входить в этот вопрос. Будем считать, скажем, координатную систему «неподвижных» звезд той, к которой мы относим все движения.

Итак, вопрос о теории электромагнитных явлений в движущихся телах коренной и сам по себе (как вопрос о движущихся телах на Земле), и в отношении влияния движения самой Земли уточняется в указанной новой фазе следующим образом. Учение об электромагнитных явлениях в покоящихся телах вылилось в руках Максвелла в систему уравнений. Как обобщить эти уравнения на случай движущихся тел?

Максвелл сам неоднократно обращался к этому вопросу, но решение его, выдержанное в духе всей теории Максвелла, облек в определенную форму Герц. На этих герцевских уравнениях, которые сыграли несомненную роль в дальнейшем развитии, мы коротко и остановимся. Я не буду при этом строго придерживаться хода мыслей Герца, а изложу идею его теории по существу.

Герц исходит из максвелловских уравнений, которые должны быть справедливы для покоящихся тел. В интегральной форме эти уравнения таковы:

$$\oint_L H_s ds = \frac{1}{c}\left\{\frac{d}{dt}\int_S D_n dS + \int_S J_n dS\right\},$$

$$\oint_L E_s ds = -\frac{1}{c}\frac{d}{dt}\int_S B_n dS.$$

В дифференциальной форме:

$$\operatorname{rot} \mathbf{H} = \frac{1}{c}\left(\frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}\right),$$

$$\operatorname{rot} \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}.$$

Без связи между **B**, **H**, **D** и **E** эти уравнения не имеют смысла. Мы ограничимся изотропными телами, для которых

$$\mathbf{D} = \varepsilon \mathbf{E}, \qquad \mathbf{B} = \mu \mathbf{H}, \qquad \mathbf{J} = \sigma \mathbf{E}.$$

Мы имеем далее

$$\operatorname{div} \mathbf{D} = \rho, \qquad \operatorname{div} \mathbf{B} = 0.$$

Кроме этих соотношений между чисто электрическими величинами, для того чтобы их можно было связать с *силами*, задается еще выражение для энергии. Ее плотность

$$w = \frac{\mathbf{ED} + \mathbf{HB}}{2}.$$

Здесь все написано в так называемых рациональных лоренцовых единицах, в которых единица заряда в $\sqrt{4\pi}$ раз меньше абсолютной электростатической единицы, а единица напряженности — в $\sqrt{4\pi}$ раз больше. Благодаря этому в большинстве формул исчезает 4π, но, конечно, по существу ничего нового в этом нет.

Все написанные выше уравнения относятся к *неподвижным телам*.

Вообще любые уравнения имеют смысл лишь тогда, когда указано, в какой системе координат они справедливы. Утверждается, что система координат, в которой с достаточной точностью справедливы написанные уравнения, *существует*, а именно, что эта система, определяемая некими тремя неподвижными точками,— система неподвижных звезд. Таким образом, написанные уравнения относятся к телам, неподвижным в этой системе. Между тем

они применялись — в сущности, непоследовательно — к движущимся телам, например к динамомашинам. Строго говоря, это недопустимо: к трансформатору их можно применять, но к динамомашине нельзя. В подобных случаях поступали примерно следующим образом. Предположим, что мы имеем катушку и магнит. Возьмем магнит в положениях 1 и 2, возьмем разность магнитных потоков и разделим на $t_2 - t_1$. Таким образом, в выражение для электродвижущей силы косвенно войдет скорость движения магнита, но специфическое влияние движения — если оно существует — при таком квазистационарном подходе не учитывается. *Приближенно* для *малых* скоростей такой подход все же допустим, так как он дает в этом случае достаточно точные результаты. Но в описанных опытах это уже не годится, и необходимо было распространить уравнения на движущиеся тела.

Герц исходил из того, что *эфир полностью увлекается телами*. Замечу сразу же: в оптике это приводит к неверным выводам, так как, по Герцу, коэффициент увлечения должен равняться единице. Таким образом, герцевские уравнения не охватывают оптических явлений. Интересно и вместе с тем странно, что Герц — второй создатель электромагнитной теории света — отказался от объяснения оптических явлений. Он сам говорит, что хочет охватить только электромагнитные явления в узком смысле. Это первый недостаток его теории: заведомо исключается целый класс явлений.

Второй недостаток — теория Герца во многих вопросах не дает однозначного ответа. Наконец, в-третьих, по отношению к описанным выше опытам она дает просто неверный ответ.

Спрашивается, зачем же тогда рассматривать теорию Герца? Прежде всего важно и интересно знать, к каким заключениям приводит предположение о полном увлечении эфира, какие случаи находятся в согласии с этим предположением. Но, кроме того, сам характер постановки вопроса у Герца очень интересен и предвосхищает его постановку в теории относительности.

Сделаем здесь небольшое отступление и сформулируем нашу задачу. Ее можно разбить на две:

1) Как выглядят в нашей неподвижной системе координат уравнения для электромагнитных явлений, протекающих в движущихся телах? Это вопрос физический.

2) Если известны уравнения в неподвижной системе координат, то как они будут выглядеть в другой системе, например в движущейся (лаборатория на Земле и т. п.)?

Второй вопрос дорелятивистская физика считала вопросом простого пересчета от одной системы к другой, вопросом чисто кинематическим. Заслуга Эйнштейна в том, что он показал, что вся трудность именно во *втором* вопросе — как перейти от одной

системы к другой. Он показал, что здесь молчаливо делались «само собой понятные», но ни на чем не основанные допущения. Это типично: долго ищут в каком-либо направлении, но вот приходит крупный человек и говорит: вы напрасно трудитесь; весь вопрос именно в том, что для всех «понятно».

Второй вопрос возникает и в других областях физики. Посмотрим, как с ним обстоит дело в механике, скажем, в механике точки. Уравнения механики — основные ньютоновские уравнения движения — справедливы в некоторой определенной системе координат (в системе неподвижных звезд)

$$m\frac{d^2\mathbf{r}}{dt^2} = \mathbf{R}$$

или в компонентах

$$m\frac{d^2x}{dt^2} = X,$$

$$m\frac{d^2y}{dt^2} = Y,$$

$$m\frac{d^2z}{dt^2} = Z.$$

Мы хотим теперь перейти к другой, движущейся системе. Однако вопрос о переходе возникает и без движения: как выглядят наши уравнения в новой системе, которая просто сдвинута и повернута по отношению к первой? Решение этого вопроса, конечно, общеизвестно, но для нас оно важно.

Введем некоторые предварительные понятия, с тем чтобы приучиться к определенному способу изложения. Прежде всего нам нужно выразить x, y, z через координаты x', y', z', относящиеся к другой системе. Как известно, это будет линейное преобразование

$$x' = a + \alpha_1 x + \beta_1 y + \gamma_1 z,$$

$$y' = b + \alpha_2 x + \beta_2 y + \gamma_2 z,$$

$$z' = c + \alpha_3 x + \beta_3 y + \gamma_3 z,$$

где α_1, β_1, γ_1 — косинусы углов между осью x' и соответственно осями x, y, z и т. д. Если обе системы — прямоугольные декартовы, то имеют место соотношения ортогональности

$$\left.\begin{array}{ll} \alpha_1\alpha_2 + \beta_1\beta_2 + \gamma_1\gamma_2 = 0, & \alpha_1^2 + \beta_1^2 + \gamma_1^2 = 1, \\ \dots\dots\dots\dots & \dots\dots\dots\dots \end{array}\right\}$$

Но в таких обозначениях все чрезвычайно громоздко. Удобнее

ввести для координат обозначения x_i и x_i' ($i = 1, 2, 3$), и тогда преобразование запишется в виде

$$x_i' = a_i + \sum_{k=1}^{3} b_{ik} x_k, \qquad i = 1, 2, 3,$$

а соотношения ортогональности — в виде

$$\sum_{k=1}^{3} b_{ik} b_{jk} = \delta_{ij} = \begin{cases} 1 & \text{при } i = j, \\ 0 & \text{при } i \neq j. \end{cases}$$

Мы воспользовались евклидовой геометрией. В этом уже лежит известное ограничение, о котором прежде никто не говорил, но которое сохраняется, правда, и в специальной теории относительности, где мы также стоим на точке зрения евклидовой геометрии.

Чтобы выполнить преобразование уравнений движения к новой системе координат, нам теперь остается написать слагающие силы в новой системе. Мы утверждаем — и это *физическое* утверждение,— что слагающие силы в новой системе выражаются через ее слагающие в старой таким же образом, как разности новых координат x_i' через разности старых x_i, т. е.

$$X_i' = \sum_{k=1}^{3} b_{ik} X_k, \qquad i = 1, 2, 3.$$

Это физическое утверждение вот в каком смысле: пусть на веревке висит груз; что такое слагающие силы веса? Представим себе оси координат в виде гладких стержней, на которые надеты пружины. По растяжению этих пружин мы и судим о величине слагающих. Тогда *оказывается*, т. е. это говорит *опыт*, что растяжения пружин в новой и старой системах связаны между собой именно так, как записано выше для X_i и X_i'.

Физическую величину, определяемую тремя числами, которые при изменении координатной системы преобразуются как разности координат, мы называем *вектором*. Таким образом, опыт говорит нам, что сила есть вектор.

ТРЕТЬЯ ЛЕКЦИЯ

(25.XII 1933 г.)

Краткое резюме. Исторический обзор (продолжение): откуда возникает требование векторности уравнений физики? Однородность и изотропность пространства. Принцип относительности положения и ориентировки. Преобразование Галилея. Инвариантность уравнений Ньютона и принцип относительности классической механики. Электродинамика Герца движущихся тел, полное увлечение эфира

Мы видели, что в круг нашего рассмотрения необходимо включить и электромагнитные явления. Мы видели, как Максвелл пришел к своему великому открытию, как его навело на это открытие совпадение (казавшееся тогда случайным) электромагнитной постоянной со скоростью света.

Самое точное значение электромагнитной постоянной, полученное из электрических измерений Роза и Дорсе, есть в настоящее время

$$(299\,770 \pm 30)\cdot 10^5 \ \textit{см/сек}.$$

Хотя это электромагнитное измерение и грубее, чем результат Майкельсона для скорости света, но нет никакого сомнения, что это одна и та же величина.

Мы рассмотрели ряд опытов, произведенных во второй половине XIX в. (опыты Роуланда, Рентгена, Эйхенвальда, Вильсона), относящихся к электромагнитным явлениям в движущихся телах. Во всех этих опытах речь идет о движении тел относительно наблюдателя, и все они дали положительный результат, т. е. получались явления, отсутствующие в том случае, когда тела покоятся. Эти опыты можно сравнить с опытом Физо над распространением света в воде, движущейся относительно наблюдателя, над увлечением света водой.

Но, кроме того, был сделан и ряд других электромагнитных опытов, имевших целью обнаружить влияние общего движения Земли.

Земля заряжена, и ее движение представляет собой ток, который должен создавать магнитное поле и, следовательно, действовать на магнетометр. Так, Рентген пытался обнаружить движение Земли, но ничего не получил. Де-Кудр искал влияние на явление индукции. Ранкин и Трутон пытались заметить с помощью моста Уитстона, влияет ли движение Земли на разветвление токов. Все эти опыты ничего не дали: неизменно оказывалось, что движение всего аппарата, всей лаборатории *как целого* не оказывает никакого действия, все происходит так, как будто Земля неподвижна.

Что же из этого вытекает? Чтобы вывести следствия, нужно иметь теорию электромагнитных явлений в движущихся телах. Мы указали, что вопрос, стоящий перед такой теорией, распадается на два:

1) Каковы уравнения для явлений в движущихся телах в основной координатной системе — системе неподвижных звезд?

2) Как воспринимает эти явления наблюдатель, находящийся в движущейся системе отсчета?

Второй вопрос считался вопросом простого пересчета, кинематики, математики. Именно здесь Эйнштейн и вскрыл основную трудность.

Этот второй вопрос не является специфичным для электромагнитных явлений, он возникает и в механике. Но в механике отпадает первый вопрос, потому что уравнения механики для движущихся тел и даны.

Нас интересует в первую очередь равномерное поступательное движение системы координат. Напишем уравнения механики в их классической форме и — для наглядности — для случая центральных сил, например сил всемирного тяготения

$$\frac{d^2x_i^\lambda}{dt^2} = \varkappa \sum_\mu m_\mu \frac{x_i^\lambda - x_i^\mu}{r_{\lambda\mu}^3},$$

где $\varkappa$ — гравитационная постоянная. Мы отступаем от исторического изложения с тем, чтобы подчеркнуть некоторые моменты, которые являются основными для теории относительности, но которые ранее молчаливо обходили.

Мы хотим посмотреть, как эти уравнения выглядят в движущейся системе координат. Но сначала посмотрим, как они преобразуются к другой, тоже *неподвижной* системе, т. е. при переносе начала и повороте осей. Из геометрии Евклида вытекает, что новые координаты будут

$$x_i' = a_i + \sum_{k=1}^{3} b_{ik}x_k, \qquad i = 1, 2, 3; \qquad \sum_{k=1}^{3} b_{ik}b_{jk} = \delta_{ij}$$

и, следовательно, для разностей координат мы имеем однородное (без a_i) линейное ортогональное преобразование. Мы ввели понятие вектора, назвав так совокупность трех величин, преобразующихся как разности координат. Является ли сила вектором — это вопрос опыта, измерения. Но в нашем случае это самоочевидно, так как при нашем преобразовании расстояние r не меняется, оно инвариантно, а в числителе силы как раз стоят разности координат. Таким образом, правая часть наших уравнений есть вектор. Левая часть, очевидно, тоже, так как там бесконечно малые

разности координат делятся на один и тот же промежуток времени dt. Значит, наши уравнения представляют собой равенство двух векторов и, если обозначить разность этих векторов через **A**, могут быть записаны в виде

$$\mathbf{A} = 0, \quad \text{или} \quad A_i = 0, \qquad i = 1, 2, 3.$$

Но легко доказать общую теорему: если в какой-либо системе отсчета вектор $\mathbf{A} = 0$, то он равен нулю и в новой системе. Действительно, по определению

$$A_i' = \sum_{k=1}^{3} b_{ik} A_k.$$

Но все A_k равны нулю, а значит,

$$A_i' = 0, \qquad i = 1, 2, 3.$$

Таким образом, в новой системе координат ньютоновы уравнения будут

$$\frac{d^2 x_k'^{\lambda}}{dt^2} = \varkappa \sum_{\mu} m_{\mu} \frac{x_k'^{\lambda} - x_k'^{\mu}}{r_{\lambda\mu}^3} .$$

Мы видим, что уравнения механики сохраняют свой вид при нашем преобразовании, что они *инвариантны* по отношению к этому преобразованию.

Что можно сказать об уравнениях Максвелла? **E** и **H** являются векторами, так как они определены через силы. В векторном анализе доказывается, что если **A** — вектор, то rot **A** также вектор. Следовательно, уравнения

$$\operatorname{rot} \mathbf{H} = \frac{1}{c}\left(\frac{\partial \mathbf{D}}{dt} + \mathbf{J}\right),$$

$$\operatorname{rot} \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{dt}$$

представляют собой векторные соотношения в указанном выше смысле и инвариантны по отношению к переносу начала и повороту осей. Впрочем, не совсем верно, что rot **A** является вектором, так как его слагающие преобразуются как разности координат лишь при переходе от правовинтовых систем к правым же или от левых к левым. При переходе же от правых к левым rot **A** меняет знак. Таким образом, написанные уравнения инвариантны, если дополнительно условиться о постоянном выборе винта системы. Но, даже ограничиваясь только преобразованием правых

систем в правые, можно поймать все ориентации осей. Для нашего же рассуждения важна только линейность, а не знак.

В векторном анализе доказывается далее, что для вектора **A** div **A** — инвариант. Таким образом, уравнения

$$\operatorname{div}\mathbf{D} = \rho, \qquad \operatorname{div}\mathbf{B} = 0$$

инвариантны.

Мы исходим из постулата, что все уравнения физики могут быть выражены в инвариантной форме. Этот постулат глубоко связан с *дорелятивистским представлением об однородности и изотропности пространства.* Что это значит?

Будут ли уравнения *одного и того же* движения иметь одинаковый вид в двух системах координат, повернутых друг относительно друга? Конечно, нет. Если в старой системе уравнение было только по оси z, то в новой может быть не только по z', но и по x' и y'. При преобразовании значения координат изменяются, так что, если подставить числа, уравнения будут *другие.* Инвариантность заключается не в этом. Она состоит в том, что если я помещу рассматриваемое тело относительно новой системы координат точно так же, как раньше относительно старой, то уравнения в новой системе будут тождественны уравнениям в старой. В этом и состоит однородность и изотропность пространства, наличие которых требует, чтобы *один и тот же опыт мог быть воспроизведен в разных координатных системах.*

Но если написаны уравнения механики, то еще вовсе не сказано, как движется тело. Сами по себе уравнения механики определяют еще не все, так как это *дифференциальные* уравнения и их решения содержат произвольные постоянные. Чтобы знать движение, надо знать эти постоянные, т. е. надо задать *начальные условия.*

Если даже так подобрать две системы координат, что правые части уравнений будут в обеих одни и те же, то все равно начальные условия *данного* движения будут в этих системах различны. Это и не удивительно. В этом, собственно говоря, суть поворота осей и переноса начала.

Но я могу *создать* в новой системе такие же начальные условия, как и в старой, и тогда все течение явлений относительно новых осей будет таким же, каким оно было относительно старых. В этом опять-таки проявляется однородность и изотропность пространства.

Мы видим, таким образом, что все будет протекать тождественно, если в новой системе все разместить и задать так же, как это было сделано в старой, т. е. создать такое же распределение сил и те же начальные условия. Вы не различите тогда,

повернута ли новая система или нет. То, что это так, это *физический* факт, этого никто не мог знать заранее. Логически могло бы быть и иначе, т. е. явления могли бы протекать по-разному в различно ориентированных системах. Но этого нет, нет *выделенной* системы отсчета, пространство *однородно* (перенос начала отсчета, выбор точки наблюдения не оказывают влияния) и *изотропно* (поворот осей также не влияет). Будь это не так, мы могли бы тогда, изучая явления в некоторой системе координат, установить, как она расположена и ориентирована, мы могли бы говорить тогда об абсолютных положении и направлении.

Что это означало бы? Ведь в самих понятиях положения и направления заложена относительность. Можно говорить об угле между *двумя* прямыми, а не об угле одной прямой. Но если бы одно направление было чем-то выделено, то имело бы смысл относить все направления к нему, раз и навсегда заданному, и называть ориентацию по отношению к такому направлению абсолютной.

Опыт говорит, что ни такого выделенного положения, ни выделенного направления в природе не существует. Как сказано, это физический факт, который мы и выражаем требованием инвариантности уравнений по отношению к переносу начала и повороту осей.

Вся классическая физика принимала этот *принцип относительности положения и ориентировки*. Он, естественно, напрашивается на обобщение. Как далеко оно может пойти, мы увидим далее. Пока же мы можем с уверенностью утверждать, что если в Ленинграде опыт дал иной результат, чем такой же точно опыт здесь, то значит, кто-то ошибся, а не объяснять это тем, что опыт произведен в Ленинграде.

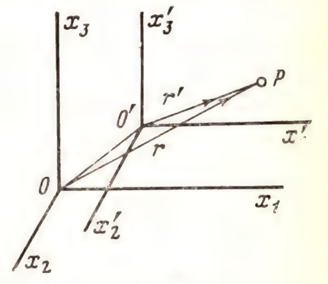


Рис. 14

Все сказанное, конечно, относится целиком и к уравнениям Максвелла.

Вернемся теперь к поставленному вопросу: как выглядят уравнения механики в новой системе координат, движущейся прямолинейно и равномерно относительно исходной системы? На основании сказанного мы можем взять оси обеих систем параллельными. Каковы же теперь уравнения преобразования? При достаточной наивности — той, которая была до принципа относительности,— их легко найти. Будем тоже наивны.

Проекция OP = проекции OO' + проекция $O'P$ (рис. 14). Пусть при $t = 0$ начала обеих систем совпадали ($OO' = O$). Тогда в момент t отрезок OO' будет иметь проекции w_1t, w_2t, w_3t, где w_i — (постоянные) компоненты скорости системы O' относи-

тельно O. Итак,

$$x_i' = x_i - w_i t, \qquad i = 1, 2, 3.$$

Дифференцируя по t, получаем

$$\frac{dx_i'}{dt} = \frac{dx_i}{dt} - w_i, \quad \text{или} \quad u_i' = u_i - w_i, \qquad i = 1, 2, 3,$$

или в векторной форме

$$\mathbf{u}' = \mathbf{u} - \mathbf{w}.$$

Это теорема сложения скоростей.

Здесь молча принимается, что, измеряя время t какого-либо события в неподвижной и движущейся системах координат, мы получаем одно и то же. Эйнштейн показал, что без дальнейшего анализа *эта фраза не имеет никакого смысла*. Но мы пока наивны и поэтому припишем к нашим уравнениям еще одно

$$t' = t.$$

Таким образом, скорость тела относительно движущейся системы есть

$$u_i' \equiv \frac{dx_i'}{dt'} = \frac{dx_i'}{dt}.$$

Дифференцируя еще раз, получаем

$$\frac{d^2x_i'}{dt'^2} = \frac{d^2x_i}{dt^2}.$$

Следовательно, в уравнениях Ньютона, содержащих *именно вторые* производные (обстоятельство, которого никто не мог знать a priori), левые части не изменяются. Что касается правых частей, то они тоже не меняются, так как зависят только от разностей координат, из которых $\mathbf{w}$ выпадает,

$$x_i'^{\lambda} - x_i'^{\mu} = x_i^{\lambda} - x_i^{\mu}.$$

Таким образом, *уравнения Ньютона инвариантны по отношению к рассматриваемому так называемому галилеевскому преобразованию*.

Опять-таки отсюда не следует, что некоторое *данное* явление наблюдается в обеих системах одинаково. Если, например, точка покоится относительно одной системы, то относительно другой она движется. Здесь опять играет роль различие начальных условий. Но если начальные условия взять в обеих системах

одинаковыми, то и движения будут одинаковы. Значит, существует бесконечное множество систем, движущихся прямолинейно и равномерно относительно исходной (и относительно друг друга), в которых законы механики одинаковы. *Нет какой-либо выделенной* системы, скажем той же системы, связанной с неподвижными звездами.

Могло бы быть и не так. Могло бы быть, например, что уравнения Ньютона справедливы только в системе неподвижных звезд. Тогда существовало бы абсолютное движение.

Разумеется, движение, понимаемое как изменение положения тел, может быть только относительным, абсолютное движение, т. е. перемещение безотносительно к чему-либо,— бессмыслица. Но если бы всегда можно было относить движение к одной определенной, выделенной системе, то имело бы смысл назвать его абсолютным. Мы знаем, что *и этого* нет. Однако последнее утверждение уже не содержится в самом понятии движения, а выражает самостоятельный физический факт.

Этот *принцип относительности равномерного поступательного движения* можно формулировать двояко:

1) уравнения Ньютона инвариантны при галилеевском преобразовании;

2) все механические явления при одинаковых начальных условиях протекают одинаково во всех системах, движущихся прямолинейно и равномерно относительно системы неподвижных звезд (так называемых *инерциальных* системах).

Сказать 2) еще не значит сказать 1), ибо второе утверждение является гораздо более общим. Классическая физика отождествляла оба утверждения, так как считала галилеевское преобразование единственно справедливым. До теории относительности обе формулировки просто сливались. *Теория же относительности, полностью сохраняя* 2), *отрицает справедливость и уравнений Ньютона и преобразования Галилея.*

То, что принцип относительности равномерного прямолинейного движения действительно является *физическим* законом, видно уж из того, что он не распространяется на ускоренно движущиеся системы. Пусть, например,

$$x_i' = x_i - g_i t^2$$

(равноускоренное движение). Тогда меняются вторые производные и вид уравнений не сохраняется. Следовательно, даже при одинаковых начальных условиях явления будут протекать в обеих системах различно.

Известно, что поступательное движение Земли не влияет на механические явления, но ее суточное вращение сказывается. Легко видеть, какую форму принимают уравнения движения

в равномерно вращающейся системе отсчета. Применив преобразование к вращающимся осям

$$\left.\begin{aligned} x' &= x\cos\omega t + y\sin\omega t,\\ y' &= -x\sin\omega t + y\cos\omega t,\\ z' &= z, \end{aligned}\right\}$$

мы получим

$$m\ddot{x}' = X' - 2m\omega\dot{y}' + m\omega^2 x',$$
$$\dots\dots\dots\dots\dots\dots$$

где

$$\left.\begin{aligned} X' &= X\cos\omega t + Y\sin\omega t,\\ &\dots\dots\dots\dots\dots \end{aligned}\right\}$$

Таким образом, уравнения совсем другие и все явления происходят иначе, чем в исходной системе. Именно поэтому мы не можем обнаружить по механическим явлениям наличие поступательного движения Земли, но можем узнать, вращается она или нет. Одним из многих опытов является маятник Фуко. Другой — отклонение падающих тел от вертикали. В 1919 г. очень тонкий опыт был сделан Гагеном (Hagen). На нити подвешены массивные шары, разведенные распоркой. Момент инерции всей системы I, а угловая скорость (из-за вращения Земли) ω. Если удалить распорку, то шары сойдутся, момент инерции уменьшится до значения I' и по закону сохранения момента ($\omega I = \omega' I'$) должно возникнуть угловое ускорение. Явление обусловлено, таким образом, вращением Земли.

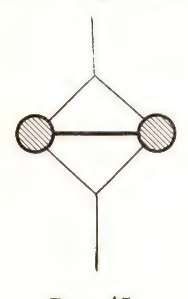


Рис. 15

Итак, механический принцип относительности справедлив только для неускоренных систем отсчета. В связи с этим еще одно замечание. Когда мы говорим, что все механические явления протекают в двух инерциальных системах одинаково, мы предполагаем, что *все* условия повторены в обеих системах.

Строго говоря, для этого обе системы должны были бы одинаково двигаться по отношению к системе неподвижных звезд. Но в такой форме принцип не имеет никакой физической ценности. Он ценен потому, что он справедлив для всякой *замкнутой* системы, т. е. такой, от которой все массы, не движущиеся вместе с ней, достаточно удалены. Интерес представляет именно то, что равномерно и прямолинейно движущуюся систему при достаточном удалении прочих масс *можно* считать замкнутой, а при вся-

ком другом движении — уже *нельзя*. Общая теория относительности уже не может отвлекаться от действия удаленных масс.

И, наконец, заметим следующее. Мы еще не знаем, каковы уравнения электродинамики для движущихся тел, в механике же твердого тела или системы точек этот вопрос отпадает. Но можно, правда несколько искусственно, поставить этот вопрос для механических явлений в сплошных средах, например для распространения звука в воздухе. Пусть мы знаем скорость звука c относительно воздуха и хотим узнать, какова скорость звука, если воздух движется, и вообще хотим написать уравнения гидродинамики для движущегося воздуха.

Мы знаем, что галилеевское преобразование сохраняет вид уравнений, и поэтому можем сразу же сказать, что скорость звука будет просто складываться со скоростью воздуха. Воздух полностью увлекает с собой звук в противоположность тому, что нашел Физо для света, где увлечение лишь частичное. По сути дела, здесь и скрыта вся трудность электродинамики движущихся тел.

Переходим теперь к электромагнитным явлениям. Здесь наш первый вопрос существует: как выглядят уравнения электродинамики для движущихся тел? Этого нельзя было вывести, это нужно было угадать, нужно было сделать новый шаг. Максвелл и Герц пытались дать уравнения для движущихся тел и, хотя сами они этого не говорят, несомненно сделали это с желанием удовлетворить принципу относительности механики, т. е. получить уравнения, инвариантные по отношению к преобразованию Галилея.

Есть одно отличие явлений электромагнитных и, конечно, оптических от, скажем, механических, например упругих, колебаний и т. п., отличие, *характерное* для фарадей-максвелловской точки зрения, не имевшее места в прежних взглядах на электрические явления, но существовавшее в прежних взглядах на оптические явления и связанное с конечной скоростью распространения света. Эта особенность та, что мы должны признать, что в пустоте есть поля. Мы *должны* говорить о полях в пустоте. Часто мы выражаем это, говоря — в *экспериментальной* пустоте.

Во всяком случае, максвелловская теория построена так, что она говорит о полях в пустоте.

Как же вводит максвелловская теория пустоту (эфир) в рассмотрение, чем она ее характеризует? Она считает эфир изолятором, принципиально ничем не отличающимся от всякого другого изолятора — стекла и т. п.,— и как каждый (изотропный) изолятор характеризует его двумя величинами и ничем более. Это диэлектрическая постоянная ε и магнитная проницаемость μ. Эфир отличается от стекла тем же, чем отличается стекло от

кварца, воды и т. п. (проводимость, наличие потерь — привходящие вещи).

Итак, эфир трактуется паритетно, ничем принципиально не выделяясь из класса других изоляторов. Это чрезвычайно типичная черта максвелловской теории.

Но при переходе к движущимся телам у нас здесь сразу возникает трудность, и трудность, принципиальная для всего максвелловского воззрения. Мы увидим далее, в чем она заключается.

Это стремление не разделять эфир и материю, стремление характеризовать все как одно целое сохранено и у Герца. Герц оставляет в силе максвелловские уравнения в их *интегральной* форме, но с одним небольшим изменением, в котором и заключается вся суть. Для неподвижных тел мы имели уравнения

$$\left.\begin{aligned} &\oint_L H_s ds - \frac{1}{c}\left\{\frac{d}{dt}\int_S D_n dS + \int_S J_n dS\right\}, \\ &\oint_L E_s ds = -\frac{1}{c}\frac{d}{dt}\int_S B_n dS, \end{aligned}\right\} \qquad (4)$$

причем контур L и опирающаяся на него поверхность S тоже покоились. Поэтому $\frac{d}{dt}\int = \int\frac{\partial}{\partial t}$, что и приводит к обычной дифференциальной форме максвелловских уравнений. Герц говорит: уравнения (4) с $\frac{d}{dt}\int\limits_S$ остаются в силе, но теперь, когда тела движутся,

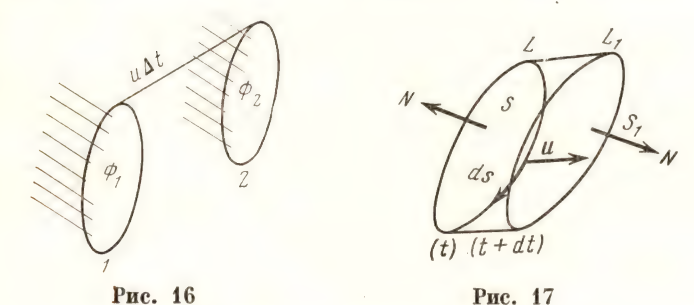


Рис. 16 Рис. 17

нужно брать поверхность S, *закрепленную в теле*, жестко связанную с частицами движущейся материи. Это существенно, ибо изменение потока через неподвижный и через движущийся контур — не одно и то же. Для движущегося контура скорость изменения потока

$$\lim_{\Delta t\to 0}\frac{\Phi_2 - \Phi_1}{\Delta t}$$

слагается из: 1) изменения Ф вследствие явной зависимости поля от времени и 2) изменения, обусловленного перемещением контура. Для неподвижного контура поток меняется только в силу первой причины, но для движущегося контура поток, вообще говоря, будет переменным, даже если поле постоянно.

Итак, контур L и поверхность S движутся вместе с материей. Дальнейшее — уже простая математика.

Мы хотим написать уравнения Герца в дифференциальной форме. Для этого надо преобразовать выражение скорости изменения потока через движущуюся поверхность. Согласно предыдущему, эта скорость изменения потока состоит из двух членов:

$$\frac{d}{dt}\int\limits_S D_n dS = \int\limits_S \frac{\partial D_n}{\partial t}\,dS + \lim_{\Delta t \to 0}\frac{\int\limits_{S_1} D_N dS + \int\limits_S D_N dS}{\Delta t}.$$

Знак «плюс» в последнем члене появляется потому, что правовинтовая нормаль $\mathbf{n}$, по которой должен вычисляться поток, направлена на S противоположно нормали $\mathbf{N}$, внешней по отношению к объему, прочерченному поверхностью S за время Δt. Обозначим через x сумму интегралов по S_1 и по S

$$x = \int\limits_{S_1} D_N dS + \int\limits_S D_N dS.$$

Тогда поток наружу через замкнутую поверхность ($S + S_1 +$ + поясок между L и L_1) равен

$$x + \Delta t \oint\limits_L [d\mathbf{s}, \mathbf{u}]\,\mathbf{D},$$

так как элемент поверхности пояска есть $\mathbf{N}\,dS = [d\mathbf{s}, \mathbf{u}]\,\Delta t$. Ввиду того что

$$\oint\limits_L [d\mathbf{s}, \mathbf{u}]\,\mathbf{D} = -\oint\limits_L [\mathbf{D}, \mathbf{u}]\,d\mathbf{s} = -\int\limits_S \operatorname{rot}_n [\mathbf{D}, \mathbf{u}]\,dS,$$

полный поток через замкнутую поверхность ($S + S_1 +$ поясок) можно записать в виде

$$x - \Delta t \int\limits_S \operatorname{rot}_n [\mathbf{D}, \mathbf{u}]\,dS.$$

С другой стороны, этот полный поток по теореме Гаусса должен быть равен интегралу от $\operatorname{div}\mathbf{D}$ по объему, прочерченному поверхностью S. Элемент этого объема есть $dS \cdot u_n \Delta t$, так что

полный поток будет

$$\Delta t \int\limits_S u_n \operatorname{div} \mathbf{D} dS.$$

Приравняв оба последних выражения, найдем x, а подставив x в выражение для $\frac{d}{dt}\int\limits_S D_n dS$, получим

$$\frac{d}{dt}\int\limits_S D_n dS = \int\limits_S \left\{\frac{\partial D_n}{\partial t} + u_n \operatorname{div} \mathbf{D} + \operatorname{rot}_n [\mathbf{D}, \mathbf{u}]\right\} dS.$$

Это выражение для скорости изменения потока, если подставить его в уравнения (4) и учесть, что поверхность S произвольна, тотчас же дает нам дифференциальную форму уравнений Герца

$$\left.\begin{aligned} \operatorname{rot} \mathbf{H} &= \frac{1}{c}\left\{\frac{\partial \mathbf{D}}{\partial t} + \operatorname{rot} [\mathbf{D}, \mathbf{u}] + \mathbf{u} \operatorname{div} \mathbf{D} + \mathbf{J}\right\}, \\ \operatorname{rot} \mathbf{E} &= -\frac{1}{c}\left\{\frac{\partial \mathbf{B}}{\partial t} + \operatorname{rot} [\mathbf{B}, \mathbf{u}]\right\}. \end{aligned}\right\} \tag{5}$$

Остальные уравнения

$$\operatorname{div} \mathbf{D} = \rho, \qquad \operatorname{div} \mathbf{B} = 0,$$

равно как и энергетические соотношения, остаются неизменными. Таковы уравнения Герца для движущихся тел, записанные в нашей неподвижной системе отсчета. Тело в этих уравнениях характеризуется величинами ε, μ и $\mathbf{u}$, а явления в каждой точке — векторами $\mathbf{E}$, $\mathbf{H}$. Если всюду $\mathbf{u} = 0$, то получаются обычные максвелловские уравнения.

Теперь второй вопрос: как относятся эти уравнения к преобразованию Галилея? Рассуждая формально, это легко установить, выполнив преобразование Галилея от x_i, t к x_i', t'. Но при этом возникает вопрос о том, как преобразуются $\mathbf{E}$ и $\mathbf{H}$, т. е. меняются или не меняются напряженности полей при переходе к другой системе отсчета. Герц утверждает, что

$$\mathbf{E}' = \mathbf{E}, \qquad \mathbf{H}' = \mathbf{H},$$

т. е. электрическое и магнитное поля абсолютны. Тогда, если произвести пересчет, оказывается, что уравнения инвариантны, т. е. сохраняют свой вид. Например, второе из уравнений (5) в движущейся системе будет

$$\operatorname{rot}' \mathbf{E}' = -\frac{1}{c}\left\{\frac{\partial \mathbf{B}'}{\partial t'} + \operatorname{rot}' [\mathbf{B}, \mathbf{u} - \mathbf{w}]\right\},$$

где rot' берется по x', y', z'; $\mathbf{w}$ — скорость этих осей по отношению к осям x, y, z, т. е.

$$\mathbf{u} - \mathbf{w} = \mathbf{u}'$$

— скорость тела относительно новых (штрихованных) осей. Точно так же сохраняют свой вид и остальные уравнения.

Этот результат можно было предвидеть заранее, уже из основных положений. Герц говорит, что поверхность движется вместе с телом. Значит, если вы движетесь вместе с телом, то относительно вас она неподвижна и, следовательно, во всякой системе, которая движется вместе с телами, справедливы обычные максвелловы уравнения.

С этой физической точки зрения можно усмотреть даже гораздо больше. Если система движется не поступательно, а вращается, но как одно целое (скажем, как Земля вместе с вами и со всеми вашими аппаратами), то относительно вас по-прежнему все неподвижно. А это означает, что в противоположность уравнениям механики уравнения Герца инвариантны относительно любого вращения и вообще любого ускоренного движения рассматриваемой системы как твердого тела. По Герцу, на электромагнитных явлениях не сказывается не только движение Земли по орбите, но и ее суточное вращение. Таким образом, уравнения Герца удовлетворяют гораздо более общему принципу относительности, чем уравнения Ньютона.

Но высшая инстанция для всякой теории — опыт. Как же теория Герца согласуется с опытом?

Опыт говорит, что электромагнитные явления не зависят от поступательного движения Земли, и в этом уравнения Герца находятся в согласии с опытом, так как они дают такую независимость, и притом в любом приближении. Однако уже здесь выявляется принципиальная трудность, присущая *всей* герцевской теории. Принцип относительности требует для одинакового протекания явлений в разных системах отсчета, чтобы *вся* конфигурация была в них воспроизведена одинаково. Но в герцевской теории скорость $\mathbf{u}$ относится не только к телам, но *и к пустоте.* В члене $\mathbf{u}\,\text{div}\,\mathbf{D} = \rho\mathbf{u}$, $\mathbf{u}$ не вызывает вопросов, так как это движение зарядов, а заряды несет только тело. Но, чтобы имел смысл член $\text{rot}\,[\mathbf{D}, \mathbf{u}]$, надо знать $\mathbf{u}$ в пустоте, ибо поле есть и в пустоте. Таким образом, для того чтобы уравнения Герца имели смысл, приходится допустить, что эфир имеет скорость и что эта скорость есть $\mathbf{u}$, т. е. эфир *полностью увлекается телами.* Иначе не выходит принцип относительности или же не выходят уравнения Максвелла.

Допущение полного увлечения эфира сразу же вступает в противоречие с явлением аберрации и с опытом Физо, который пока-

зывает, что увлечение не полное, а частичное (по Френелю). Таким образом, уравнения Герца формально удовлетворяют принципу относительности, но это согласие только кажущееся. Оно требует таких утверждений о движении эфира, которые противоречат аберрации и коэффициенту увлечения.

Опыты Роуланда, Рентгена, Эйхенвальда, Вильсона мы рассмотрим в следующий раз и увидим, что они указывают на это же слабое место теории Герца — молчаливо допускаемое полное увлечение. Наконец, и то следствие теории Герца, что вращение Земли не должно влиять на электромагнитные явления, тоже опровергается опытом Майкельсона.

ЧЕТВЕРТАЯ ЛЕКЦИЯ

(1.I 1934 г.)

Исторический обзор (продолжение): инвариантность уравнений Герца при преобразовании Галилея, теория Герца и опыт, опыт Саньяка и опыт Майкельсона, явление индукции, опыт Физо, электромагнитные опыты, необходимость отказа от полного увлечения эфира

Мы рассмотрели принцип относительности положения и ориентировки, выражающий представление об однородности и изотропности пространства. Затем мы перешли к другому кругу вопросов, связанных с равномерным и прямолинейным движением друг относительно друга систем, ориентировка которых сохраняется при этом движении неизменной. Здесь мы имеем *принцип относительности механики*, который мы дали в двух формулировках. Первая, чисто математическая, говорит о том, что ньютоновы уравнения инвариантны по отношению к преобразованию Галилея: $x_i' = x_i - w_i t$, $t' = t$. Вторая, гораздо более общая, опирается на предположение, что преобразование Галилея характеризует движение одной системы относительно другой, отображает физический процесс. Тогда принцип относительности механики означает, что существует бесчисленное множество систем, движущихся друг относительно друга прямолинейно и равномерно, в которых при одинаковых начальных условиях все механические явления протекают одинаково. Это уже физическое утверждение. Следствием его является классический закон сложения скоростей.

Мы утверждаем, таким образом, что во всех указанных систе-

мах, в том числе и в системе неподвижных звезд, одинаковые механические опыты дают одни и те же результаты. Следовательно, равномерное поступательное движение системы нельзя обнаружить никакими механическими опытами внутри этой системы. Для вращающихся систем это уже не так. На них принцип относительности не распространяется, и было бы неверно утверждать, что в одной вращающейся системе явления будут протекать так же, как и в других.

Откуда берется самое название «принцип относительности»? Если бы существовала какая-либо выделенная система, скажем, такая, в которой справедливы законы Ньютона, в то время как в других справедливы иные законы, то тогда имело бы смысл считать ее неподвижной и говорить о скорости по отношению к ней как об *абсолютной* скорости. Принцип относительности утверждает, что такой системы *нет* и что поэтому можно говорить только об *относительных* скоростях. Насколько удачен сам термин «принцип относительности» — это вопрос другой.

Мы перешли затем к электромагнитным явлениям в движущихся телах и рассмотрели то решение вопроса об уравнениях для этих тел, которое было дано Герцем. Вместо максвелловских уравнений для неподвижных тел Герц пишет уравнения

$$\mathrm{rot}\,\mathbf{H} = \frac{1}{c}\left\{\frac{\partial \mathbf{D}}{\partial t} + \mathbf{u}\,\mathrm{div}\,\mathbf{D} + \mathrm{rot}\,[\mathbf{D}, \mathbf{u}] + \mathbf{J}\right\},$$

$$\mathrm{rot}\,\mathbf{E} = -\frac{1}{c}\left\{\frac{\partial \mathbf{B}}{\partial t} + \mathrm{rot}\,[\mathbf{B}, \mathbf{u}]\right\}.$$

При этом существенно следующее. Для того чтобы эти уравнения имели смысл, каждый их член должен иметь смысл в каждой точке. Но поле существует и там, где нет тел. Между тем члены $\mathrm{rot}\,[\mathbf{D}, \mathbf{u}]$, $\mathrm{rot}\,[\mathbf{B}, \mathbf{u}]$ там не исчезают. Что же представляет собой $\mathbf{u}$ в этих точках, что надо понимать под скоростью там, где нет тел? Эта трудность, коренящаяся в том, что у Максвелла вакуум ничем принципиально не отличается от тел, губит всю герцеву электродинамику по существу.

Мы убедились далее, что уравнения Герца инвариантны при преобразовании:

$$x_i' = x_i - w_i t, \qquad t' = t, \qquad u_i' = u_i - w_i,$$

$$\mathbf{E}' = \mathbf{E}, \qquad \mathbf{H}' = \mathbf{H},$$

откуда следует, что в новой (штрихованной) системе все электромагнитные явления протекают так же, как и в неподвижной.

В частности, если я двигаюсь вместе с телами и приборами, которые все имеют скорость $\mathbf{u}$ относительно исходной неподвижной системы, то для меня относительная скорость тел $\mathbf{u}'$ будет

равна нулю. Но для моей движущейся системы отсчета справедливы те же уравнения Герца

$$\operatorname{rot}' \mathbf{H}' = \frac{1}{c}\left\{\frac{\partial \mathbf{D}'}{\partial t} + \mathbf{u}' \operatorname{div}' \mathbf{D}' + \operatorname{rot}' [\mathbf{D}', \mathbf{u}'] + \mathbf{J}'\right\},$$

$$\operatorname{rot}' \mathbf{E}' = -\frac{1}{c}\left\{\frac{\partial \mathbf{B}'}{\partial t} + \operatorname{rot}' [\mathbf{B}', \mathbf{u}']\right\}.$$

Полагая в них $\mathbf{u}' = 0$, мы получаем просто уравнения Максвелла. Таким образом, начав с определенной координатной системы — системы неподвижных звезд,— мы видим, что можно исходить из любой галилеевой системы, что и в отношении электромагнитных явлений все равномерно и прямолинейно движущиеся системы эквивалентны друг другу.

Но относительная скорость $\mathbf{u}'$ должна быть равна нулю *всюду*, т. е. эфир должен двигаться с той же скоростью, что и тела, должен полностью увлекаться как внутри, так и вне тел. Следовательно, Герц объясняет независимость электромагнитных явлений от поступательного движения Земли, но этот успех теории покупается ценой полного увлечения эфира.

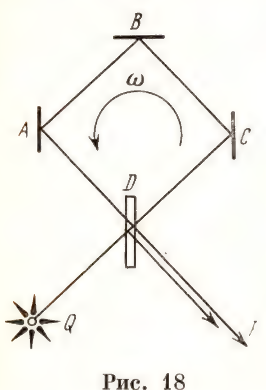


Рис. 18

Инвариантность уравнений Герца при преобразовании Галилея ясна из самого вида этих уравнений в интегральной форме. В системе, движущейся вместе с телами, сказать, что поверхность S жестко связана с частицами материи, это значит сказать, что она неподвижна относительно координатных осей. Но если поверхность неподвижна относительно этих осей, то мы имеем уравнения Максвелла. Это простое рассуждение сразу же показывает нам, что мы можем не ограничиваться прямолинейным и равномерным движением. В самом деле, допустите, что вы движетесь со всеми телами и аппаратурой *ускоренно*. Поверхность, жестко закрепленная в теле, все равно будет относительно вас неподвижна, если только тела не испытывают деформации. Следовательно, уравнения Герца инвариантны относительно *всякого* движения системы как твердого тела [1]. Что же говорит опыт?

В 1912 г. Гаррес, а затем Саньяк произвели так называемый вихревой опыт, с тем чтобы установить, увлекается ли эфир при вращении системы. Эта система представляла собой устройство

[1] [Это связано, как нетрудно понять, с тем, что в уравнения Герца, в отличие от уравнений Ньютона, входят только первые полные производные по времени.]

из трех зеркал A, B, C и полупрозрачной пластинки D, смонтированных на площадке, которую можно было приводить во вращение вместе с источником света Q и фотопластинкой в P. Луч света Q расщепляется пластинкой D на два луча — $DABCDP$ и $DCBADP$. Эти лучи, обегающие контур $ABCD$ во встречных направлениях, интерферируют в P, и *если эфир не увлекается*, то при вращении системы должно получиться смещение полос.

Действительно, в отсутствие вращения время, затрачиваемое светом на обход замкнутого пути $ABCD$, равно

$$T = \oint \frac{ds}{c}.$$

Если система вращается с угловой скоростью $\vec{\omega}$, то скорость света относительно системы будет равна c за вычетом слагающей линейной скорости вдоль пути света в данной точке этого пути, т. е. будет равна

$$c_1 = c - u_s = c - [\vec{\omega}, \mathbf{r}]\, \mathbf{s}_1,$$

где $\mathbf{s}_1$ — единичный вектор в направлении распространения света в данной точке. Ограничиваясь величинами первого порядка относительно $\frac{\omega r}{c}$, получаем, что время обхода $ABCD$ будет теперь

$$T_1 = \oint \frac{ds}{c_1} \approx \oint \frac{ds}{c} + \frac{1}{c^2} \oint [\vec{\omega}, \mathbf{r}]\, \mathbf{s}_1 \cdot dS = T + \frac{1}{c^2} \oint [\mathbf{r}, d\mathbf{s}]\, \vec{\omega}.$$

Но интеграл $\oint [\mathbf{r}, d\mathbf{s}]$ равен просто $2S\mathbf{n}$, где S — площадь, охватываемая контуром пути света, а $\mathbf{n}$ — правовинтовая нормаль к плоскости этого контура. Таким образом,

$$T_1 = T + \frac{2S\omega}{c^2} \cos\left(\widehat{\mathbf{n}, \vec{\omega}}\right).$$

Для второго луча, обегающего контур по часовой стрелке, нормаль направлена противоположно $\mathbf{n}$, так что для него

$$T_2 = T - \frac{2S\omega}{c^2} \cos\left(\widehat{\mathbf{n}, \vec{\omega}}\right).$$

Следовательно, при неувлекаемом эфире уже в первом порядке должно получиться запаздывание одного луча относительно другого, равное

$$\Delta T = T_1 - T_2 = \frac{4S\omega}{c^2} \cos\left(\widehat{\mathbf{n}, \vec{\omega}}\right),$$

и интерференционные полосы должны сместиться. По Герцу же, т. е. при полностью увлекаемом эфире, не должно быть никакого эффекта. Саньяк наблюдал смещение полос, прекрасно согласующееся с выведенной формулой.

Естественно возникает вопрос о том, как скажется вращение самой Земли. Угловая скорость здесь мала, и это сильно затрудняет опыт, так как приходится брать большие площади S. Такой опыт, являющийся аналогом механического опыта Фуко с маятником, удалось осуществить Майкельсону в 1925 г. Это повторение опыта Саньяка, но вращение установки обусловлено вращением Земли. Так как менять ω при этом невозможно, то Майкельсон сделал два контура — большой $DABCD$, с площадью около 208 тыс. *м*^{2}, и малый $DAB'C'D$ (рис. 19). Первоначально вследствие тепловых потоков в воздухе интерференционная картина получилась очень неустойчивой. Поэтому Майкельсон проложил вдоль всего пути света около 2 *км* труб, из которых воздух был выкачан. Он получил тогда смещение полос на долю $0{,}230 \pm 0{,}005$ полосы, в то время как теория давала 0,236. Итак, Герц не прав.

Здесь есть, правда, одна тонкость, на которую, быть может, стоит указать. Разумеется, какое-то количество воздуха в трубах все-таки оставалось, а тогда совершенно ясно, чтó означает скорость $\mathbf{u}'$, и так как уравнения инвариантны, то ответ теории Герца неверен. Но если выкачать из труб абсолютно все, то Герц

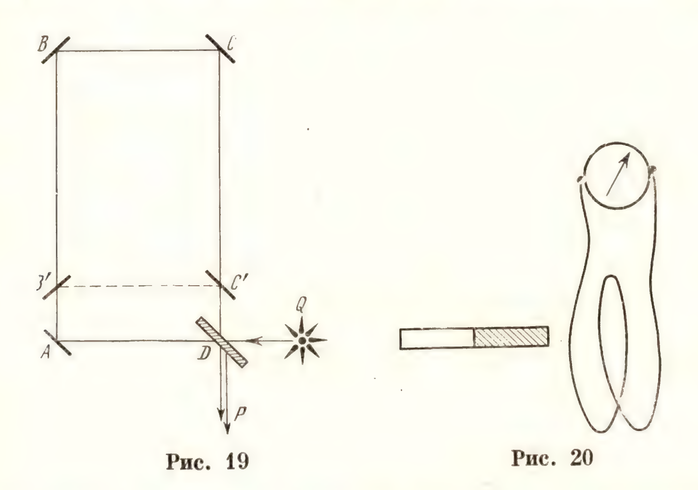


Рис. 19 Рис. 20

мог бы сказать, что он не знает, чтó следует подставлять в качестве $\mathbf{u}'$ в пустоте. Может быть, там, где ничего нет, эфир и не увлекается. Легко видеть, однако, чтó при такой постановке вопроса возникает трудность другого рода. Получается, что наличие сколь угодно малого количества воздуха дает полное увлечение эфира, а неувлечение наступает лишь в абсолютной пустоте. Физически ясно, что переход к пустоте должен быть непрерывным. Мы не можем допустить такого нелепого положения, когда присутствие

минимального количества воздуха само по себе не оказывает никакого влияния (коэффициент преломления практически равен 1), но *движение* этого воздуха влияет самым существенным образом, обусловливая *полное* увлечение эфира.

Посмотрим теперь, что говорит теория Герца в тех случаях, когда имеются относительные движения тел, т. е. тела движутся относительно друг друга и относительно наблюдателя внутри лаборатории. Существует множество такого рода опытов, и мы рассмотрим теперь некоторые из них с точки зрения теории Герца.

Возьмем, например, явление индукции. Опыт говорит, что здесь важно только *относительное* движение магнита и катушки (рис. 20). Теория Герца, очевидно, согласуется с этим, так как в ней все определяется как раз относительным движением. Тем не менее есть известная разница в подходе к этим двум случаям. Если в нашей координатной системе движется магнит, то мы имеем переменное магнитное поле, которое обусловливает возникновение вихревого электрического поля. Последнее и создает электродвижущую силу в неподвижной катушке. Таким образом, здесь индукция вызывается членом $\frac{\partial \mathbf{B}}{\partial t}$ в уравнении для $\operatorname{rot} \mathbf{E}$. Если же магнит неподвижен в нашей системе, а движется катушка, то никакого электрического поля вообще нет и то, что индукционный ток все-таки возникает, обусловлено теперь наличием у катушки скорости $\mathbf{u}$, т. е. членом $\operatorname{rot}\,[\mathbf{B}, \mathbf{u}]$ в уравнении для $\operatorname{rot} \mathbf{E}$. Имеется, таким образом, хотя и не наблюдаемое, но принципиальное различие в толковании этих двух случаев.

Опыт Физо. Нас интересует, какова скорость света относительно неподвижного наблюдателя (система x, y), если свет распространяется в воде, движущейся со скоростью w. Для наблюдателя, движущегося вместе с водой (система x', y'), уравнения Герца переходят в уравнения Максвелла и скорость света будет $c_1 = c/n$. Для неподвижного же наблюдателя скорость света c_1' согласно галилееву преобразованию просто сложится из c_1 и скорости воды w, т. е. будет $c_1' = c_1 + w$. Однако на опыте получается $c_1' = c_1 + w\,(1 - n^{-2})$, т. е. входит коэффициент увлечения, которого не знает теория Герца.

Представим себе, что опыт сделан не с водой, а с воздухом, который мы делаем все более и более разреженным. При этом $n \to 1$ и коэффициент увлечения непрерывно стремится к нулю. По Герцу же, как бы ни был воздух разрежен, своим движением он всегда меняет скорость на w, т. е. $c' = c + w$ $(n \approx 1)$, и лишь в пустоте $(n = 1)$ скачком должно получиться $c' = c$. В макроскопической теории такой переход к пределу неудовлетворителен. Были попытки обойти эту трудность ссылкой на аналогич-

ные явления в теории газов; как известно, вязкость и теплопроводность газа не зависят от давления, так что и здесь мы переходим к вакууму без постепенного уменьшения вязкости и теплопроводности. Но это просто неверно. Теория, из которой вытекает независимость вязкости и теплопроводности от давления, применима только к неразреженным газам. Когда средняя длина свободного пробега молекул становится сравнима с размерами сосуда, то эта теория уже неверна и сами явления вязкости и теплопроводности вообще пропадают. Здесь же, в электромагнетизме, само явление (распространение света) остается и в вакууме, так что трудность с отсутствием непрерывного перехода к вакууму не устраняется.

Но Герц с самого начала отказался от объяснения оптических явлений, так что мы известным образом не вправе требовать от его теории таких объяснений. Вернемся к электромагнитным опытам в узком смысле.

Опыт Роуланда с вращением заряженного металлического диска. Вращение здесь, конечно, ни при чем, и для понимания идеи опыта можно взять поступательное движение заряженной обкладки. Мы имеем уравнение

$$\operatorname{rot} \mathbf{H} = \frac{1}{c}\left\{\frac{\partial \mathbf{D}}{\partial t} + \mathbf{u} \operatorname{div} \mathbf{D} + \operatorname{rot}\,[\mathbf{D}, \mathbf{u}] + \mathbf{J}\right\},$$

из которого следует, что движущийся заряд ($\mathbf{u} \operatorname{div} \mathbf{D} = \rho\mathbf{u}$) в магнитном отношении эквивалентен току проводимости $\mathbf{J}$. Член

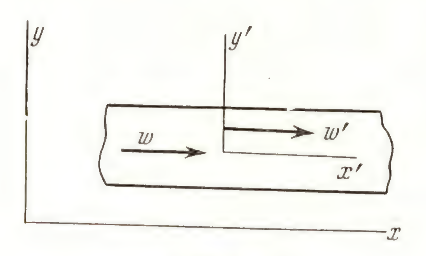


Рис. 21

$\frac{\partial \mathbf{D}}{\partial t} = 0$, так как поле стационарно, а член $\operatorname{rot}\,[\mathbf{D}, \mathbf{u}] = 0$, так как диэлектрик между разноименными обкладками неподвижен, а в эбоните нет поля. Наше уравнение выведено в предположении непрерывности всех величин. Но в опыте Роуланда целесообразно говорить о поверхностной плотности заряда на проводнике $\sigma = \operatorname{Div} \mathbf{D}$ (Div — поверхностная дивергенция) и о поверхностном токе $\sigma\mathbf{u}$. Согласно Герцу, вращение обкладок должно

давать такое же магнитное поле, какое было бы при неподвижных обкладках, если по ним пропустить ток $i = \sigma u = \varepsilon u E$ (так как в воздушном зазоре между + и — обкладкой $\mathrm{Div}\,\mathbf{D} = D = \varepsilon E$). Роуланд подтвердил этот вывод, а позднее Эйхенвальд проверил его количественно с большой точностью. Таким образом, для конвекционного тока теория Герца дает правильный результат.

Опыты Рентгена и Эйхенвальда. В этих опытах между неподвижными обкладками движется диэлектрик, и поэтому в уравнении (благодаря стационарности поля, неподвижности обкладок и отсутствию тока проводимости) остается в правой части только член $\mathrm{rot}\,[\mathbf{D}, \mathbf{u}]$. Можно было бы, конечно, рассматривать непрерывное убывание $\mathbf{u}$ к краям диэлектрика, но фактически он двигался весь как целое, так что и здесь целесообразно перейти к разрывной трактовке, введя поверхностный ротор $\mathbf{g} = \mathrm{Rot}\,\mathbf{A} = [\mathbf{n}, \mathbf{A}_1 - \mathbf{A}_2]$, где $\mathbf{A}_1$ и $\mathbf{A}_2$ — значения вектора по обе стороны от поверхности, нормаль к которой есть $\mathbf{n}$. Учитывая, что вне диэлектрика $\mathbf{u} = 0$, что $\mathbf{n}$ перпендикулярно $\mathbf{u}$ и параллельно (или антипараллельно) $\mathbf{D}$, мы получаем, что поверхностный ротор $[\mathbf{D}, \mathbf{u}]$, т. е. плотность эквивалентного поверхностного тока, по абсолютной величине есть

$$i = \varepsilon\,|\,uE\,|.$$

Эйхенвальд получил экспериментально

$$i = (\varepsilon - 1)\,|\,uE\,|.$$

Таким образом, опять в теории Герца получается абсурдный результат, что при $\varepsilon = 1$, т. е. при переходе к вакууму, ток не пропадает.

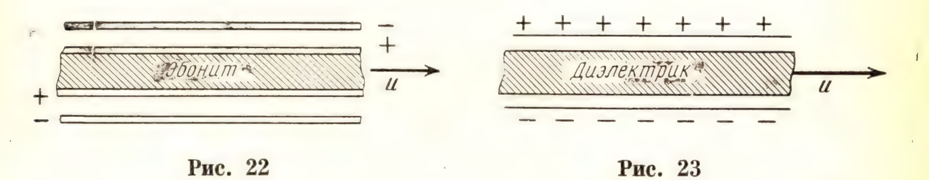


Рис. 22 Рис. 23

Если в формуле Герца умножить скорость u на $\frac{\varepsilon - 1}{\varepsilon}$, то получится правильное выражение для i. Другими словами, надо вместо скорости u ввести $u' = u\,(1 - n^{-2})$, т. е. ввести коэффициент увлечения. Конечно, это формальное рассуждение, но не лишено интереса, что можно сохранить уравнения Герца, считая, что при движении диэлектрика играет роль не вся скорость, а лишь часть ее, соответствующая коэффициенту увлечения.

Эйхенвальд сделал еще и такой опыт, когда все движется вместе — и диэлектрик, и обкладки. По Герцу, в этом случае сумма токов Роуланда и Рентгена равна нулю, т. е. не должно быть никакого магнитного поля. Эйхенвальд же получил эффект, не зависящий от ε диэлектрика и подтверждающий тем самым наличие частичного увлечения. Почему для конвекционного тока теория Герца дает правильный результат, а для тока Рентгена — неверный? Потому что движение зарядов всегда связано с движением материи, между тем как движение поля, по воззрениям Максвелла — Герца, не обязательно связано с движением материального диэлектрика [1].

Опыт Вильсона также можно схематизировать, заменив вращение поступательным движением. Диэлектрик движется между обкладками и пронизывается магнитным полем, перпендикулярным к плоскости рисунка. Ввиду постоянства магнитного поля мы имеем $\frac{\partial \mathbf{B}}{\partial t} = 0$ и, следовательно, $\mathbf{E}$ удовлетворяет уравнению $\operatorname{rot}\left(\mathbf{E} + \frac{1}{c}[\mathbf{B}, \mathbf{u}]\right) = 0$, откуда с точностью до градиента произвольной функции, который мы можем приравнять нулю, $\mathbf{E} = -\frac{1}{c}[\mathbf{B}, \mathbf{u}]$. Так как $\mathbf{E} = -\operatorname{grad}\varphi$, то, интегрируя это равенство от одной обкладки до другой, получаем

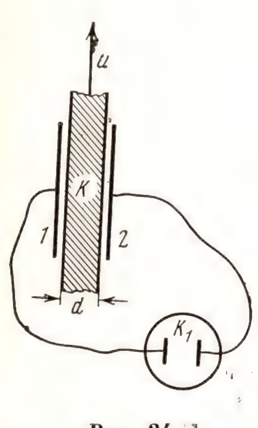


Рис. 24

$$\int_1^2 \mathbf{E} ds = \varphi_1 - \varphi_2 = \frac{|Bu|d}{c}.$$

Если учесть теперь, что наш конденсатор с емкостью K шунтируется электрометром, емкость которого равна K_1, то измеряемая разность потенциалов V составит, очевидно, долю $K/(K+K_1)$ от $\varphi_1 - \varphi_2$

$$V = \frac{|Bu|d}{c} \frac{K}{K+K_1}.$$

Опыт дал иной результат, а именно:

$$V = \left(1 - \frac{1}{\varepsilon}\right) \frac{|Bu|d}{c} \frac{K}{K+K_1},$$

т. е. внес опять ту же поправку $(1 - n^{-2})$.

---

[1] [См. примечание 1 на стр. 90.]

Мы видим, что недостаток теории Герца всегда в одном и том же — в полном увлечении эфира. Не только оптические, но и собственно электромагнитные опыты говорят о том, что полного увлечения нет. Но это полное увлечение вытекает из требования, чтобы в электродинамике выполнялся принцип относительности. Значит, надо отказаться от принципа относительности в электродинамике. На этот путь стал Лоренц.

ПЯТАЯ ЛЕКЦИЯ

(7.I 1934 г.)

Исторический обзор (продолжение): силы в электродинамике Герца, какова должна быть правильная теория, практический принцип относительности, электронная теория Лоренца, электромагнитный импульс и момент импульса, усреднение уравнений Лоренца для медленно движущихся тел, вопрос о выполнении практического принципа относительности в теории Лоренца

Мы разобрали теорию Герца. В общем, дело обстояло так: уравнения Максвелла — Герца для движущихся тел инвариантны по отношению к преобразованию Галилея, т. е. они сохраняют свой вид во всех равномерно поступательно движущихся системах, а так как считается, что это преобразование отображает пространственно-временные соотношения в движущихся телах, то принцип относительности в теории Максвелла — Герца соблюден.

Громадное число опытов показывает, что поступательное движение Земли действительно не сказывается на электромагнитных явлениях, и здесь, таким образом, все как будто бы обстоит хорошо. Но теория Герца идет дальше: ее уравнения инвариантны при всяком движении системы как твердого тела. Однако имеется целый ряд опытов с телами, движущимися относительно Земли (опыты Физо, Эйхенвальда, Вильсона и др.), которые не подтверждают теории Герца, не объясняются ею совсем или же объясняются только качественно, а количественно получается типичное для теории расхождение (отсутствие коэффициента увлечения). В итоге теория Герца неудовлетворительна, не отвечает тем требованиям, которые надо предъявлять к физической теории.

Есть еще один принципиальный вопрос, на который Герц определенным образом отвечает, но этот ответ снова заставляет

задуматься. Речь идет о силах, с которыми электромагнитное поле действует на тела. Рассмотрим для простоты только случай твердых тел, т. е. оставим в стороне деформации. Принципиально это упрощение ничего не меняет.

В электродинамике Максвелла — Герца силы, действующие в поле на тело, всегда могут быть сведены к натяжениям, т. е. к *поверхностным* силам плотности

$$\mathbf{T} = \mathbf{E}(\varepsilon\mathbf{E},\ \mathbf{n}) - \frac{\mathbf{n}\varepsilon E^2}{2} + \mathbf{H}(\mu\mathbf{H},\ \mathbf{n}) - \frac{\mathbf{n}\mu H^2}{2}.$$

Уравнения движения какой-либо системы тел, заключенных в объеме V, ограниченном поверхностью S, имеют при этом вид

$$\frac{d}{dt}\sum m\mathbf{v} = \oint_S \mathbf{T}dS,$$

$$\frac{d}{dt}\sum [\mathbf{r},\ m\mathbf{v}] = \oint_S [\mathbf{r},\ \mathbf{T}]\,dS.$$

Отсюда можно сделать следующее заключение: если поверхность S настолько удалена, что на ней поле равно нулю (замкнутая система), то правые части исчезают, т. е. равнодействующие всех сил и всех моментов сил, действующих на замкнутую систему тел, равны нулю. Иначе можно сказать, что в замкнутой системе тел результирующее количество движения и результирующий момент количества движения постоянны:

$$\sum m\mathbf{v} = \text{const}, \qquad \sum [\mathbf{r},\ m\mathbf{v}] = \text{const}.$$

Это является выражением III закона Ньютона — равенства действия и противодействия. Таким образом, герцевские силы гарантируют выполнение III закона Ньютона, и с этой точки зрения они могут нас удовлетворить. Но если рассчитать соответствующие объемные силы, то плотность объемных сил будет

$$\mathbf{f} = \mathbf{f}_1 + \mathbf{f}_2,$$

где $\mathbf{f}_1$ — сила, которая при наличии тел отлична от нуля как в переменном, так и в постоянном поле, но которая исчезает в отсутствие вещества. Сила же $\mathbf{f}_2$ имеет следующий вид:

$$\mathbf{f}_2 = \frac{\varepsilon\mu}{c^2}\frac{\partial\mathbf{S}}{\partial t},$$

где $\mathbf{S} = c\,[\mathbf{E}, \mathbf{H}]$ — вектор Пойнтинга. Таким образом, $\mathfrak{f}_2$ не исчезает и в пустоте, если только поле переменно, т. е., по Герцу, силы действуют и на эфир. Только благодаря этому в переменных полях соблюдается закон равенства действия и противодействия. То, что, по Герцу, на эфир должны действовать силы, следует из конечной скорости распространения. Пусть, например, плоская поверхность, помещенная в вакууме, односторонне излучает плоскую электромагнитную волну (рис. 25). Благодаря реакции поверхность начнет двигаться, и если центр тяжести остается в покое, то только потому, что такая же сила приложена в поле волны к эфиру. Итак, выполнение III закона Ньютона покупается ценой действия сил на эфир. Но никому и никогда не удавалось обнаружить это действие, заставить эфир двигаться, хотя делалось множество таких опытов (например, прохождение света около быстро вращающегося диска). Герц исходил из того, что эфир — такая же среда, как и всякое другое вещество, и здесь этот взгляд мстит за себя. Нет проверки этого, нет реального содержания. Все это, равно как и то обстоятельство, что сколь угодно разреженное тело должно полностью увлекать эфир, ставит концепцию Герца под подозрение. Наконец, прямое противоречие с опытом окончательно заставляет от нее отказаться.

Между прочим, Ф. Франк показал, что требования, чтобы уравнения для равномерно и прямолинейно движущихся тел 1) обращались для покоящихся тел в уравнения Максвелла и 2) были инвариантны по отношению к преобразованию Галилея, однозначно приводят к уравнениям Герца. Поэтому закрадывается сомнение: можно ли вообще удовлетворить принципу относительности в электродинамике? Это сомнение было уже у Герца; он кончает свою статью словами, что «правильная теория, вероятно, будет различать движение тел и движение эфира». Именно по этом пути пошел Лоренц.

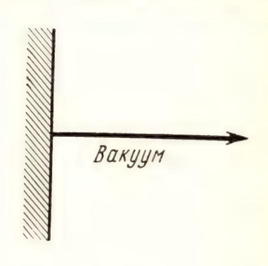


Рис. 25

Чего же нужно требовать от правильной теории?

Мы заранее знаем, что она не будет удовлетворять принципу относительности: этому учит уже опыт Физо. Для неподвижного наблюдателя скорость света в покоящейся воде равна c_1, а в движущейся $c_1 + w\,(1 - n^{-2})$. Это говорит опыт. Принцип же относительности требует, чтобы для наблюдателя, движущегося вместе с водой, скорость света была c_1. Но если галилеево преобразование верно, если оно правильно дает переход от неподвижной

системы к движущейся, то для наблюдателя, движущегося вместе с водой, скорость будет

$$[c_1 + w(1 - n^{-2})] - w = c_1 - wn^{-2},$$

а не c_1, т. е. принцип относительности неверен. Можно было бы возразить, что для применимости принципа относительности надо воспроизвести одинаковые условия также и в отношении источника света, т. е. при движении воды (для того чтобы скорость света относительно нее была c_1) надо заставить и источник двигаться вместе с водой. Однако это не спасает дела, ибо мы исходим из *независимости скорости света от движения источника.* Если бы такая зависимость была, подобно тому, как в механике скорость снаряда складывается со скоростью пушки, то все могло бы быть в порядке. Но опыты, как мы увидим далее, полностью подтверждают, что скорость света не зависит от движения источника. Исходя из этого факта, мы должны заключить, что либо галилеево преобразование верно и тогда в электродинамике не может быть сохранен принцип относительности, либо принцип относительности остается в силе, но тогда надо отказаться от галилеева преобразования. На этот последний шаг еще никто не мог осмелиться. Сохраняя же преобразование Галилея, мы вынуждены оставить принцип относительности. Но ведь опыты говорят, что движение Земли не сказывается на скорости света! Что значит «не сказывается»?

Посмотрим еще раз, что говорят опыты.

Движение Земли не влияет на электрические и оптические явления... С какой точностью? Ведь до сих пор мы рассматривали только такие опыты, которые были рассчитаны на обнаружение эффектов первого порядка относительно w/c. Вот с *этой* точностью мы знаем, что движение Земли не влияет. Правда, опыт Физо дает отличие уже в первом порядке, но по отношению к движению Земли такого опыта не делали. Мы, конечно, *можем* такой опыт сделать, но сначала сделайте, а тогда будем о нем говорить. Далее, еще никогда не делались опыты с двумя часами (эти опыты теория относительности вообще квалифицирует особо). Все опыты, фактически сделанные, относились к сравнению пути *двух* лучей, и все делались в *первом* порядке w/c.

Вот с этой оговоркой, что не все опыты делались, а те, которые делались, имели точность до первого порядка, поступательное движение Земли не сказывается на электромагнитных явлениях. Назовем это «практическим принципом относительности». Опыт Физо показывает, что, сохраняя галилеево преобразование, мы не можем удовлетворить точному принципу относительности (независимость электромагнитных явлений в любых опытах, в любом порядке), но теория *должна объяснить практический*

принцип. Можно ли построить *такую* теорию? Эту задачу поставил и разрешил Лоренц.

К этому времени появился целый ряд новых экспериментальных фактов. Интерес был привлечен к явлениям электрического разряда в газах, к катодным лучам, которые были открыты Плюкером еще в 1859 г., но природа которых была установлена именно в это время. Герц считал их продольными электромагнитными волнами. Оказалось, что это движущиеся заряженные тельца, масса которых зависит от скорости, причем они имеют различные продольную и поперечную массу. Помимо объяснения практического принципа относительности, теория Лоренца должна была охватить огромный круг вопросов, и она блестяще с этим справилась. Она сумела объяснить множество явлений и в неподвижных телах (дисперсию и абсорбцию, зееман-эффект и т. д.), но основное в теории Лоренца — это движущиеся тела, для которых она и была главным образом развита. Фундаментальная работа Лоренца так и называется «Versuch einer Theorie der elektrischen und magnetischen Erscheinungen in bewegten Körpern». Именно эта сторона теории Лоренца — явления в движущихся телах — будет интересовать и нас.

Лоренц постулирует существование всепроникающего, однородного, изотропного и *неподвижного* эфира. Покоящаяся относительно эфира система отсчета — это особая, выделенная система. Таким образом, Лоренц *сознательно отказывается от принципа относительности.*

Эфир у Лоренца — единственный диэлектрик в максвелловском смысле слова, тела же представляют собой совокупности положительных и отрицательных заряженных частиц. Отрицательные частицы — электроны — все одинаковые по своей массе (когда они покоятся относительно эфира) и по величине заряда. Это именно те частицы, поток которых образует катодные «лучи». Положительные частицы всегда связаны с «весомой» материей, с веществом. Свойства тел различны в силу разного их построения из этих положительных и отрицательных частиц. Таким образом, в мире нет ничего, кроме этих зарядов и эфира. Проводники — это тела, в которых электроны в промежутках между соударениями могут свободно двигаться. В диэлектриках каждая молекула нейтральна, причем разноименные заряды связаны в ней какими-то силами, которые в первом приближении можно считать квазиупругими. Кроме того, электроны могут вращаться вокруг собственных осей и совершать орбитальные движения в молекулах, чем объясняются магнитные свойства тел.

Теперь, с точки зрения квантовой теории, все это нам кажется, конечно, немного наивным и крайне упрощенным, но эта ясная и простая картина дала чрезвычайно много. Надо заметить, что

она куда более стройна, чем современная квантовая теория (конечно, надо учесть, что последняя еще слишком нова), но сейчас она, разумеется, может рассматриваться только как модель.

Мы отвлечемся от магнитных свойств, ибо пара- и диамагнетики ведут себя почти как вакуум (говоря по Максвеллу, $\mu = 1$), а поведения ферромагнетиков теория Лоренца вообще объяснить не может. Точно так же мы исключаем из рассмотрения не так давно открытые так называемые сегнетоэлектрики, аналогичные ферромагнетикам, но не по своим магнитным, а по электрическим свойствам.

Пусть $\mathbf{e}$ и $\mathbf{h}$ — напряженности электрического и магнитного полей, характеризующие состояние чистого эфира, вакуума. Для них Лоренц сохраняет максвелловские уравнения для пустоты

$$\operatorname{rot}\mathbf{h} = \frac{1}{c}\frac{\partial \mathbf{e}}{\partial t}, \qquad \operatorname{rot}\mathbf{e} = -\frac{1}{c}\frac{\partial \mathbf{h}}{\partial t},$$
$$\operatorname{div}\mathbf{e} = 0, \qquad \operatorname{div}\mathbf{h} = 0.$$

Как же связаны $\mathbf{e}$ и $\mathbf{h}$ с частицами, с зарядами? Здесь по существу имеется два вопроса: 1) как поле создается зарядами и 2) как поле действует на заряды. На первый вопрос Лоренц отвечает следующим образом. Там, где плотность заряда $\rho \neq 0$,

$$\operatorname{div}\mathbf{e} = \rho.$$

Далее, из опытов Роуланда Лоренц заимствует эквивалентность движущегося заряда току и полагает поэтому

$$\operatorname{rot}\mathbf{h} = \frac{1}{c}\left(\frac{\partial \mathbf{e}}{\partial t} + \rho\mathbf{v}\right).$$

Написанные уравнения говорят о том, как поле создается движущимися зарядами. Второй же вопрос — о действии поля на движущиеся заряды — решается тем, что на основании закона Био — Савара Лоренц постулирует следующее выражение для плотности силы:

$$\mathbf{f} = \rho\left(\mathbf{e} + \frac{1}{c}[\mathbf{v}, \mathbf{h}]\right).$$

В совокупности приведенные уравнения должны описать решительно *все*, включая и явления в движущихся телах, ибо движение в них уже включено. В этом смысле они аналогичны уравнениям Ньютона. Лоренцовы уравнения сохраняются и в теории относительности, с ними работают и поныне.

Необходимо заметить следующее. Во-первых, концепция, согласно которой вся материя (вернее, проводники, ибо до Максвелла об электрических свойствах диэлектриков вообще знали

мало) построена из электрических зарядов, движущихся в вакууме, не была новой. Из нее исходили уже теории Вебера, Римана, Клаузиуса. Но от их воззрений лоренцову теорию принципиально отличает то, что взаимодействие частиц происходит по Максвеллу, т. е. действие распространяется в поле с *конечной* скоростью, в то время как в старых теориях принималось мгновенное дальнодействие.

Во-вторых, в противоположность Максвеллу — Герцу, Лоренц принципиально *отказывается от действия сил на эфир*: сила пропорциональна ρ и там, где $\rho = 0$, она обращается в нуль. Этот отказ очень важен и влечет за собой глубокие следствия.

Плотность $\mathfrak{f}$ лоренцовой силы можно преобразовать к виду

$$\mathfrak{f} = \mathfrak{f}^{\text{H}} - \frac{1}{c^2}\frac{\partial \mathbf{S}}{\partial t}, \tag{6}$$

где S — вектор Пойнтинга, а $\mathfrak{f}^{\text{H}}$ — плотность силы, по Герцу. Таким образом, уничтожается как раз тот член, который давал у Герца силу, действующую на эфир. Из (6) следует, что результирующая сила $\mathbf{F}$ и результирующий момент $\mathbf{N}$, действующие на заряды, заключенные в объеме V, ограниченном поверхностью S, равны соответственно

$$\mathbf{F} = \int_V \mathfrak{f}\,dV = \oint_S \mathbf{T}\,dS - \frac{d}{dt}\int_V \frac{\mathbf{S}}{c^2}\,dV,$$

$$\mathbf{N} = \int_V [\mathbf{r}, \mathfrak{f}]\,dV = \oint_S [\mathbf{r}\,\mathbf{T}]\,dS - \frac{d}{dt}\int_V \frac{[\mathbf{r}, \mathbf{S}]}{c^2}\,dV.$$

Введем полное количество движения $\mathbf{G}_0$ и полный момент количества движения $\mathbf{Y}_0$ тел, находящихся в нашем объеме V:

$$\mathbf{G}_0 = \sum m\mathbf{v}, \qquad \mathbf{Y}_0 = \sum [\mathbf{r}, m\mathbf{v}].$$

Тогда уравнения движения для всей системы зарядов в объеме V будут

$$\frac{d\mathbf{G}_0}{dt} = \mathbf{F} = \oint_S \mathbf{T}\,dS - \frac{d}{dt}\int_V \frac{\mathbf{S}}{c^2}\,dV,$$

$$\frac{d\mathbf{Y}_0}{dt} = \mathbf{N} = \oint_S [\mathbf{r}, \mathbf{T}]\,dS - \frac{d}{dt}\int_V \frac{[\mathbf{r}, \mathbf{S}]}{c^2}\,dV.$$

Если теперь мы поставим вопрос о полной силе, действующей на *замкнутую* систему $\left(\oint_S = 0\right)$, то, поскольку объемные инте-

гралы, вообще говоря, отличны от нуля, мы получим, что сумма всех сил отлична от нуля, т. е. закон равенства действия и противодействия не выполнен. Односторонне излучающая плоскость начнет двигаться под действием реакции, и компенсации не будет.

Итак, лоренцова теория ведет к тому, что III закон Ньютона, по крайней мере в обычной своей форме, не имеет места. Центр масс системы может ускоряться без наличия внешних сил, тело само может «вытащить себя за волосы». Иначе можно сказать, что $\mathbf{G}_0$ и $\mathbf{Y}_0$ в замкнутой системе не постоянны.

Абрагам указал, однако, на то, что можно так обобщить понятия количества движения и его момента, чтобы по отношению к обобщенным величинам имел место закон сохранения. Если ввести электромагнитный импульс

$$\mathbf{G} = \int\limits_V \frac{\mathbf{S}}{c^2}\, dV,$$

локализованный в эфире, с объемной плотностью

$$\mathbf{g} = \frac{\mathbf{S}}{c^2}$$

и момент электромагнитного импульса

$$\mathbf{Y} = \int\limits_V [\mathbf{r},\ \mathbf{g}]\, dV,$$

то для замкнутой системы

$$\frac{d\mathbf{G}_0}{dt} = -\frac{d\mathbf{G}}{dt}, \qquad \frac{d\mathbf{Y}_0}{dt} = -\frac{d\mathbf{Y}}{dt}.$$

Принимая $\mathbf{G}$ и $\mathbf{Y}$ всерьез, мы получаем обобщенные законы сохранения импульса и момента импульса,— сохраняется лишь сумма соответствующих механических и электромагнитных величин

$$\mathbf{G}_0 + \mathbf{G} = \text{const}, \qquad \mathbf{Y}_0 + \mathbf{Y} = \text{const}.$$

Мы видим, что вектор Пойнтинга приобретает, помимо своего энергетического смысла, еще значение плотности импульса. Но независимо от этого толкования полученный результат имеет ряд интересных следствий (давление электромагнитных волн, отдача при излучении и т. д.). Интеграл $\int S_n dS$ дает нам поток энергии однозначно лишь в случае замкнутой поверхности. Для незамкну-

той поверхности $\int S_n dS$ уже не выражает потока энергии и лишь иногда дает приближенно правильную величину. В других случаях такое употребление **S** приводит к парадоксам. Пока **S** толкуется только как плотность потока энергии, выражение $\mathbf{S} = c\,[\mathbf{e}, \mathbf{h}]$ определено лишь с точностью до ротора произвольного вектора, так как

$$\oint\limits_S \mathrm{rot}_n\, \mathbf{a} \cdot dS \equiv 0.$$

Ничто не обязывает нас считать, что именно $c\,[\mathbf{e}, \mathbf{h}]$, а не $c\,[\mathbf{e}, \mathbf{h}] + \mathrm{rot}\,\mathbf{a}$ есть плотность потока в данной точке. Но если **S** выражает плотность электромагнитного импульса, то здесь для такой неоднозначности уже нет места.

Рассмотрим такой пример: возьмем цилиндрический конденсатор и направим по его оси однородное магнитное поле. Легко видеть, что поток **S** через всякую замкнутую поверхность равен нулю, но момент электромагнитного импульса отличен от нуля. Следовательно, если мы будем заряжать этот конденсатор, когда он уже находится в магнитном поле (до зарядки $\mathbf{Y}_0 = 0$, $\mathbf{Y} = 0$), то, поскольку $\mathbf{Y}_0 + \mathbf{Y} = 0$, а мы создаем $\mathbf{Y} = \int\limits_V \left[\mathbf{r}, \frac{\mathbf{S}}{c^2}\right] dV$, должен будет появиться механический момент количества движения $\mathbf{Y}_0 = -\mathbf{Y}$, т. е. конденсатор должен закрутиться. Наблюдать этот эффект на опыте очень трудно, так как он чрезвычайно мал, но принципиально он наблюдаем и, следовательно, позволяет определить величину **S**.

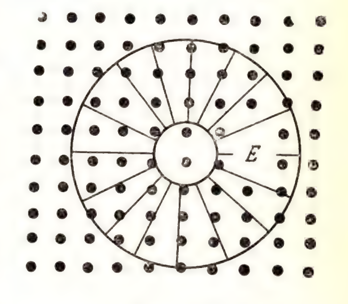


Рис. 26

Конечно, здесь ясно, в чем дело: если мы осуществляем зарядку тем, что тянем электроны по радиусу, то они заворачивают в магнитном поле и при остановке отдают свой импульс (его тангенциальная слагающая и вызывает вращение). Однако наше толкование обобщает законы сохранения, а нас интересует сейчас именно эта сторона дела. Аналогичное обобщение закона сохранения энергии (в замкнутой системе постоянна сумма механической и электромагнитной энергий) для нас привычно. Почему же не сделать аналогичное обобщение для импульса и его момента? Мы это сделали и ввели электромагнитные импульс и момент импульса. Мы увидим в дальнейшем, что в теории относительности эти обобщения вообще не независимы, а составляют один шаг.

Уравнения Лоренца написаны для истинного, микроскопического поля. Но мы не можем знать, как движутся все элементарные заряды, да это и не нужно: мы всегда измеряем средние значения, работаем со средними величинами. Как же напишутся наши электромагнитные уравнения для макроскопических, усредненных величин? При усреднении надо будет учесть строение вещества, отразить основные черты структуры проводников и диэлектриков (от магнетиков мы отвлекаемся). Мы рассмотрим сразу же общий случай движущихся тел. Это значит, что средняя скорость всех зарядов, составляющих тело, отлична от нуля. Если мы обозначим ее $\mathbf{w}$, то истинная скорость заряда будет слагаться из средней скорости тела $\mathbf{w}$ и из скорости заряда относительно тела $\mathbf{u}$

$$\mathbf{v} = \mathbf{w} + \mathbf{u}.$$

Обозначим плотность заряда электронов проводимости через ρ_l, а *среднюю* плотность их заряда $\hat{\rho}_l$ ($\hat{}$ будет обозначать усреднение) через ρ : $\hat{\rho}_l = \rho$, причем ρ — это максвелловские свободные заряды. Максвелловская плотность тока проводимости по определению $\hat{\rho}_l \mathbf{u} = \mathbf{J}$. Если имеются, кроме свободных зарядов, еще и связанные (рис. 27), то момент одной молекулы будет

$$\mathbf{p} = \sum_i e_i \mathbf{r}_i,$$

где $\mathbf{r}_i$ — расстояние заряда e_i от произвольно выбранного центра нейтральной молекулы. Средний момент единицы объема или поляризация есть

$$\mathbf{P} = N\mathbf{p},$$

где N — концентрация зарядов с моментом $\mathbf{p}$. Легко показать, что средняя плотность ρ_p связанных (поляризационных) зарядов будет

$$\hat{\rho}_p = -\operatorname{div} \mathbf{P}.$$

Таким образом, средняя объемная плотность связанных зарядов отлична от нуля только в неоднородном поле. Введем далее для усредненных напряженностей обозначения

$$\hat{\mathbf{e}} = \mathbf{E}, \qquad \hat{\mathbf{h}} = \mathbf{H}.$$

И, наконец, нам необходима связь $\mathbf{P}$ с напряженностями полей. Лоренц предполагает, что в диполях диэлектрика действуют квазиупругие силы. Тогда эта связь линейна и $\mathbf{P}$ пропорционально средней лоренцовой силе $\mathbf{E} + \frac{1}{c}[\mathbf{w}, \mathbf{H}]$. Коэффициент пропор-

циональности мы возьмем в виде $\varepsilon - 1$, где ε зависит от предположений о строении материи. Можно считать вообще, что все уравнение

$$\mathbf{P} = (\varepsilon - 1)\left(\mathbf{E} + \frac{1}{c}[\mathbf{w}, \mathbf{H}]\right)$$

является просто одним из постулатов, касающихся строения материи. Мы увидим далее, почему выбор коэффициента сделан именно в таком виде $(\varepsilon - 1)$.

Усреднение лоренцовых уравнений дает

$$\operatorname{rot}\mathbf{H} = \frac{1}{c}\left(\frac{\partial \mathbf{E}}{\partial t} + \frac{\partial \mathbf{P}}{\partial t} + \mathbf{J} + \rho\mathbf{w} + \operatorname{rot}[\mathbf{P}, \mathbf{w}]\right),$$

$$\operatorname{rot}\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{dt},$$

$$\operatorname{div}\mathbf{H} = 0, \qquad \operatorname{div}\mathbf{E} = \rho - \operatorname{div}\mathbf{P}.$$

Мы должны требовать от этих уравнений, чтобы для неподвижных тел ($\mathbf{w} = 0$) они совпадали с максвелловскими, т. е. принимали вид

$$\operatorname{rot}\mathbf{H} = \frac{1}{c}\frac{\partial \varepsilon\mathbf{E}}{\partial t}, \qquad \operatorname{rot}\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t},$$

$$\operatorname{div}\mathbf{H} = 0, \qquad \operatorname{div}\varepsilon\mathbf{E} = \rho.$$

Очевидно, при нашем выборе $\mathbf{P}$ это выполнено, так как при $\mathbf{w} = 0$ мы имеем $\mathbf{P} = (\varepsilon - 1)\mathbf{E}$, и, таким образом, под ε надо понимать просто максвелловскую диэлектрическую постоянную.

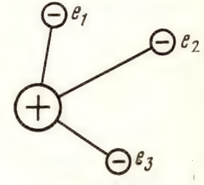


Рис. 27

Но ведь мы хотели получить больше, мы хотели объяснить абсорбцию, дисперсию и т. п. — явления, которых не объясняет теория Максвелла. Почему этого нет? Потому, что мы предположили, что отклонение зарядов в молекулах диэлектрика от их положений равновесия просто пропорционально действующей силе. Мы не учли, таким образом, массы связанных электронов, пренебрегли инерцией в уравнении движения связанного электрона. В быстропеременных полях это недопустимо, и если бы мы этого не сделали, то мы получили бы ε как функцию частоты, т. е. дисперсию.

Перейдем теперь к тому, что нас интересует в первую очередь, — к движущимся телам. Как здесь обстоит дело с толкованием электромагнитных опытов в движущихся телах? Возьмем, например,

опыт Рентгена. У Герца мы имели для плотности тока Рентгена

$$\mathrm{rot}\,[\mathbf{D},\ \mathbf{w}] = \mathrm{rot}\,[\varepsilon\mathbf{E},\ \mathbf{w}].$$

Здесь же имеем

$$\mathrm{rot}\,[\mathbf{P},\ \mathbf{w}] \approx \mathrm{rot}\,[(\varepsilon - 1)\,\mathbf{E},\ \mathbf{w}]$$

(с точностью до членов первого порядка относительно $\mathbf{w}$). Таким образом, в согласии с результатами Эйхенвальда плотность тока Рентгена пропорциональна $\varepsilon - 1$, а не ε. Мы указали при рассмотрении теории Герца, что ее можно приспособить к опытам, если формально взять вместо $\mathbf{w}$ только часть скорости, а именно увлекаемую (френелевскую) часть. Здесь это отпадает. По Лоренцу, дело не в $\mathbf{w}$, а в $\mathbf{P}$. $\mathbf{w}$ — полная скорость материи, но эфир неувлекаем, а движутся только заряды, и это дает именно то, что нужно. Таким образом, Лоренц совершенно естественно объясняет этот опыт, равно как и все остальные электромагнитные опыты. В частности, опыт Вильсона объясняется чрезвычайно просто и наглядно тем, что поляризация движущегося в магнитном поле диэлектрика обусловливается силой $\frac{(\varepsilon - 1)}{c}[\mathbf{w},\ \mathbf{H}]$.

Теперь — самое интересное: как теория Лоренца осуществляет практический принцип относительности? Мы знаем, что точный принцип относительности заведомо не будет иметь места. Если мы, желая ответить на вопрос о том, как будут протекать электромагнитные явления для наблюдателя, движущегося вместе с телами, применим к усредненным лоренцовым уравнениям преобразование Галилея

$$x_i' = x_i - w_i t, \quad t' = t,$$

с добавлением условий

$$\mathbf{E}' = \mathbf{E}, \qquad \mathbf{H}' = \mathbf{H},$$

т. е. будем считать напряженности абсолютными, то уравнения окажутся неинвариантными. Это дает прямой расчет, но это можно усмотреть хотя бы из того, что лоренцовы уравнения отличны от герцевских уравнений, которые инвариантны.

Можно было бы пойти по пути признания неинвариантности уравнений и ограничиться показом того, что в *опытах* имеет место инвариантность. Но Лоренц, хотя он и не был математиком, к счастью, пошел другим путем. Считая, как и все, что физический смысл имеет только галилеевское преобразование, он поставил вопрос о формальных, лишенных физического содержания преобразованиях, по отношению к которым его уравнения были бы инвариантны. В качестве первого шага он берет преоб-

разование

$$x_i' = x_i - w_i t, \qquad t' = t - \frac{(\mathbf{w}, \mathbf{x}')}{c^2},$$

формально вводя новую переменную t'. В теории Лоренца она не имеет физического смысла и является только вспомогательной величиной. Физический смысл по-прежнему имеет только t, и толковать результаты надо будет в переменных $\mathbf{x}'$ и t.

Самое интересное то, что по отношению и к этому преобразованию лоренцовы уравнения тоже неинвариантны...

ШЕСТАЯ ЛЕКЦИЯ

(13.I 1934 г.)

Краткое резюме. Исторический обзор (продолжение): метод преобразований, объяснение опыта Физо и частичного увлечения эфира в теории Лоренца, опыт Майкельсона, опыт Трутона и Нобля, гипотеза сокращения, развитие метода преобразований, необходимость отказа от преобразования Галилея

Мы установили общее положение: если мы хотим сохранить и независимость скорости света от движения источника и преобразование Галилея и в то же время хотим остаться в согласии с опытом, то нужно отказаться от принципа относительности. На этот путь и стал Лоренц. Он принимает, что существует однородная, изотропная и неувлекаемая среда — эфир. Система отсчета, связанная с этой средой, является *выделенной* системой, и движение по отношению к ней имеет смысл называть *абсолютным*. Таким образом, принцип относительности нарушен.

Мы рассмотрели силы, действующие, согласно Лоренцу, на материальные тела, и пришли к выводу, что закон равенства действия и противодействия для них не выполнен. Силу оказалось возможным разложить на две слагающие: силу герцевского типа, удовлетворяющую закону равенства действия и противодействия, и силу, не удовлетворяющую этому закону. Мы получили уравнения движения, которые в случае замкнутой системы имеют вид

$$\frac{d\mathbf{G}_0}{dt} = \mathbf{F} = -\frac{d}{dt}\int\limits_V \mathbf{g}\, dV, \qquad \frac{d\mathbf{Y}_0}{dt} = \mathbf{N} = -\frac{d}{dt}\int\limits_V [\mathbf{r}, \mathbf{g}]\, dV,$$

где $\mathbf{g} = \mathbf{S}/c^2$, а $\mathbf{r}$ — радиус-вектор, проведенный из фиксированной точки эфира — начала нашей системы координат — в непод-

вижную же точку наблюдения. Таким образом, полные *механические* импульс и момент импульса непостоянны, но если ввести новые понятия — *электромагнитный* импульс $\mathbf{G} = \int\limits_V \mathbf{g}\,dV$ и момент импульса $\mathbf{Y} = \int\limits_V [\mathbf{r}, \mathbf{g}]\,dV$, — то для сумм $\mathbf{G}_0 + \mathbf{G}$ и $\mathbf{Y}_0 + \mathbf{Y}$ в замкнутой системе имеет место сохранение. Отсюда вытекает ряд следствий, которые, конечно, можно получить и независимо от указанного толкования $\mathbf{G}$ и $\mathbf{Y}$. Замечу, что все эти вопросы приобретают полную ясность и точность именно в теории относительности. Здесь же все это выглядит несколько грубо.

Рассмотрим сейчас, не вдаваясь в детали и тонкости, одно следствие, по поводу которого не всегда бывает полная ясность. Пусть плоская пластинка излучила короткий электромагнитный цуг или импульс, так что поле содержится в слое толщины l (рис. 28). Этот импульс излучен в течение времени l/c и имеет энергию

$$W = \frac{l}{c} Sa = \frac{SV}{c},$$

где $V = la$ — объем слоя. Отсюда $S = \frac{cW}{V}$, и, следовательно, электромагнитный импульс излученной волны равен

$$G = \frac{S}{c^2} V = \frac{W}{c}.$$

В процессе излучения пластинка получила механический импульс, направленный в обратную сторону, т. е. $|G_0| = |Mv| = \frac{W}{c}$; следовательно, скорость, приобретенная пластинкой, будет

$$|v| = \frac{W}{Mc}.$$

Представим себе теперь, что такой импульс излучила стенка A, являющаяся дном твердого цилиндра (рис. 29). Весь цилиндр придет при этом в движение. Если стенка B полностью поглотит электромагнитную волну (спустя время L/c), то цилиндр остановится, передвинувшись на расстояние $x = v\frac{L}{c}$, причем $Mv = -\frac{W}{c}$, так что $x = -\frac{WL}{Mc^2}$.

Таким образом, x — новая координата центра тяжести. Механика считала, что внутренние силы не могут переместить центр тяжести системы. Здесь это получилось потому, что, кроме механического, мы имеем еще электромагнитный импульс. Мы оста-

немся в согласии с обычными воззрениями механики, если припишем электромагнитной энергии массу

$$m = \frac{W}{c^2}.$$

Тогда центр тяжести всей системы (цилиндр + электромагнитная энергия) останется на месте.

Такое наделение электромагнитной энергии массой получается, как вы видите, совершенно естественно. Оно напрашивается само собой, хотя оно и не вытекает ниоткуда с необходимостью. Дело здесь обстоит так же, как и с законом сохранения энергии. Когда мы заряжаем конденсатор от электростатической машины, механическая энергия исчезает. Мы говорим: она не исчезла, а превратилась в электрическую энергию конденсатора. И здесь по отношению к импульсу мы проводим тот же ход идей. Здесь нет еще никакой теории относительности. Наоборот, в релятивистской механике способ рассуждения, приведенный в нашем примере, как раз и неприменим, так как теория относительности отрицает существование абсолютно твердых тел. У нас же скорость приобретается цилиндром одновременно всем сразу, т. е. деформация распространяется быстрее c. Однако результат, полученный нами, верен и весьма важен, хотя рассуждение было и неточно: мы не учли, например, того, что стенка B движется навстречу электромагнитному импульсу. В теории относительности все эти рассуждения становятся на твердую почву.

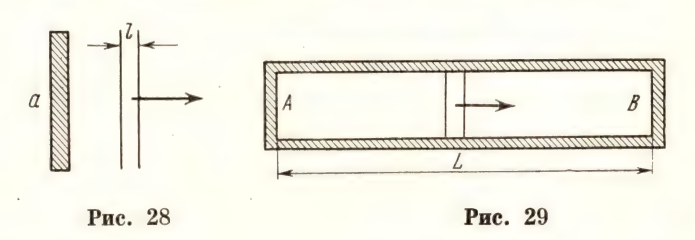


Рис. 28 Рис. 29

Мы пришли к следующим уравнениям для движущихся тел:

$$\left.\begin{aligned}
\operatorname{rot}\mathbf{H} &= \frac{1}{c}\left(\frac{\partial\mathbf{E}}{\partial t} + \frac{\partial\mathbf{P}}{\partial t} + \mathbf{J} + \rho\mathbf{w} + \operatorname{rot}[\mathbf{P},\ \mathbf{w}]\right), \\
\operatorname{rot}\mathbf{E} &= -\frac{1}{c}\frac{\partial\mathbf{H}}{\partial t}, \\
\operatorname{div}\mathbf{E} &= \rho - \operatorname{div}\mathbf{P}, \qquad \operatorname{div}\mathbf{H} = 0, \\
\mathbf{P} &= (\varepsilon - 1)\left(\mathbf{E} + \frac{1}{c}[\mathbf{w},\ \mathbf{H}]\right),
\end{aligned}\right\} \qquad (7)$$

где w, по предположению, — постоянная средняя (по всем зарядам) скорость. Согласно последнему из этих уравнений поляризация предполагается пропорциональной действующей силе, чем исключается объяснение дисперсии, магнитного вращения плоскости поляризации и т. п. Для неподвижных тел мы получаем максвелловы уравнения, чего, вообще говоря, нельзя было предвидеть из концепции Лоренца заранее. Мы видели, что с помощью лоренцовых уравнений хорошо объясняются электромагнитные опыты в движущихся телах. Далее мы поставили вопрос о том, как выполняется практический принцип относительности. Если произвести преобразование Галилея, то мы увидим, что уравнения не сохраняют своего вида. Можно было бы ограничиться констатацией этой неинвариантности уравнений по отношению к галилееву преобразованию и для каждого конкретного опыта убеждаться в том, что с достаточной точностью (до первого порядка) теория его объясняет. Но Лоренц пошел другим путем. Почему? Трудно сказать.

Лоренц ввел преобразование

$$\mathbf{r}' = \mathbf{r} - \mathbf{w}t, \qquad t' = t - \frac{(\mathbf{w}, \mathbf{r}')}{c^2}, \tag{8}$$

после которого уравнения получаются ближе по виду к максвелловским, чем при преобразовании Галилея, при условии, что напряженности **E** и **H** заменяются следующими величинами:

$$\mathbf{E}' = \mathbf{E} + \frac{1}{c}[\mathbf{w}, \mathbf{H}], \qquad \mathbf{H}' = \mathbf{H} + \frac{1}{c}[\mathbf{E}, \mathbf{w}]. \tag{9}$$

Физически нас интересуют **E** и **H** в функции $\mathbf{r}'$ и t, но мы, очевидно, сразу это получим, если найдем $\mathbf{E}'$, $\mathbf{H}'$ в функции переменных $\mathbf{r}'$, t'.

Как же выглядят преобразованные уравнения для $\mathbf{E}'$ и $\mathbf{H}'$ в новых переменных $\mathbf{r}'$ и t'? После преобразования получаем с *точностью до первого порядка относительно w/c*

$$\left.\begin{aligned} \operatorname{rot}'\mathbf{H} &= \frac{1}{c}\left(\frac{\partial \varepsilon \mathbf{E}'}{\partial t'} + \mathbf{J}\right), & \operatorname{div}'\mathbf{H}' &= 0, \\ \operatorname{rot}'\mathbf{E}' &= -\frac{1}{c}\frac{\partial \mathbf{H}'}{\partial t'}, & \operatorname{div}'\varepsilon\mathbf{E}' &= \rho - \frac{(\mathbf{J}, \mathbf{w})}{c^2}. \end{aligned}\right\} \tag{10}$$

Эти уравнения очень похожи на максвелловские, но существенно отличается последнее. В отсутствие токов проводимости ($\mathbf{J} = 0$) уравнения имеют вполне максвелловский вид.

Не совсем приятным является то, что мы пренебрегли малыми членами ($\sim w^2/c^2$) в самих дифференциальных *уравнениях*, ибо решение мало отличных уравнений может оказаться совершенно

другим. Но этот вопрос, равно как и вопросы непрерывности, мы оставим в стороне.

Интересно отметить, что Лоренц называет t' «местным временем» (t' — функция точки). В популярных книгах часто пишут, что новое время ввел именно Лоренц. Я думаю, что это не так. Из всего хода мыслей Лоренца ясно, что он считал временем t, а не t', которое было лишь вспомогательной математической величиной. С откровенностью большого человека Лоренц сам писал в 1915 г., что его ошибка именно в том и заключалась, что он считал временем t, а не t'. Это нисколько не умаляет его, а, наоборот, показывает силу его физической интуиции.

Преобразованием (8) и (9) мы пока что упростили математическое рассмотрение. В частности, в отсутствие токов проводимости для штрихованных величин все происходит просто по Максвеллу. Решения же уравнений Максвелла известны для очень большого числа задач, т. е. в ряде случаев мы имеем готовые выражения $\mathbf{E}'$ и $\mathbf{H}'$ через t' и $\mathbf{r}'$. После этого мы сейчас же можем перейти к физическим величинам $\mathbf{E}$ и $\mathbf{H}$, выраженным через физические отсчеты пространства ($\mathbf{r}'$) и времени (t).

Возьмем, например, заряженный шар. В системе координат ($\mathbf{r}'$, t')

$$\mathbf{E}' = \frac{e\mathbf{r}'}{r'^3}, \qquad \mathbf{H}' = 0.$$

Но это не физика, а математика. Нас интересуют $\mathbf{E}$ и $\mathbf{H}$. Из (9) мы видим, что $\mathbf{H}$ будет первого порядка относительно w/c, а, значит, в пренебрежении вторым порядком $\mathbf{E}' = \mathbf{E}$. Таким образом,

$$\mathbf{E} = \frac{e\mathbf{r}'}{r'^3}, \qquad \mathbf{H} = \frac{1}{c}[\mathbf{w}, \mathbf{E}] = \frac{e}{cr'^3}[\mathbf{w}, \mathbf{r}'],$$

т. е. в первом порядке шар просто переносит свое электрическое поле вместе с собой и создает (тоже движущееся вместе с ним) магнитное поле (конвекционный ток Роуланда). Уже из этого примера видна польза, приносимая методом преобразований.

Посмотрим теперь, как, с точки зрения Лоренца, объясняется опыт Физо. Пусть w — скорость воды и свет идет в том же направлении, в каком движется вода. Здесь опыть мы имеем случай диэлектрика, т. е. $\mathrm{J} = 0$. $\mathbf{E}'$ и $\mathbf{H}'$ удовлетворяют просто волновому уравнению и, следовательно, пропорциональны

$$\sin\nu\left(t' - \frac{x'}{c_1}\right), \quad \text{где } c_1 = \frac{c}{\sqrt{\varepsilon}}.$$

Чтобы судить о скорости света относительно наблюдателя, движущегося вместе с водой, надо перейти к физическим величинам, к истинному времени t и к $\mathbf{E}$, $\mathbf{H}$. Но последние являются одно-

родными линейными функциями **E**′ и **H**′ и, следовательно, тоже будут пропорциональны тому же синусу. Переходя к t, мы в первом приближении получаем

$$\sin\nu\left(t-\frac{wx'}{c^2}-\frac{x'}{c'}\right)\approx\sin\nu\left(t-\frac{x'}{c_1\left(1-\frac{c_1w}{c^2}\right)}\right),$$

т. е. скорость относительно воды есть

$$c_1'=c_1-wn^{-2},$$

в полном согласии с френелевским выражением. С точки зрения неподвижного наблюдателя, скорость будет

$$c_1'+w=c_1+w\,(1-n^{-2})$$

— классическая формула с коэффициентом увлечения. Мы можем сделать отсюда важное заключение. Теория Лоренца дает френелевский коэффициент увлечения. Но мы знаем, что наличия этого коэффициента достаточно для того, чтобы практически все оптические опыты (фактически производившиеся в первом порядке) протекали так же, как в неподвижной системе. Мы видели, что коэффициент увлечения объясняет относительность и электромагнитных опытов. Таким образом, практический принцип относительности, по крайней мере в отсутствие токов проводимости, справедлив.

Я хочу указать на одно обстоятельство, может быть и не имеющее значения, но все же не лишенное интереса. Френель предполагал эфир сгущенным в телах и говорил, что увлекается лишь сгущенная часть эфира, а несгущенная остается в покое. Все видели в этом какую-то неясность у Френеля. Об этом говорит, например, Лармор. Что значит «не увлекается та часть эфира, которая была бы, если бы не было сгущения»? В дальнейшем Стокс строит свое объяснение, просто исходя из того, что скорость эфира в теле меньше, чем снаружи (относительно тела). Эта «неясность» у Френеля приобретает в теории Лоренца вполне определенное содержание: движутся заряды, а эфир остается неподвижным. Видимо, у Френеля было предчувствие того, что что́-то движется, а что́-то остается в покое.

Возьмем теперь случай, когда есть ток проводимости. Будут ли электрические явления независимы от движения Земли и теперь? Из уравнений это непосредственно не видно, так как в них есть член, зависящий от w/c в первом порядке.

Рассмотрим, например, такой опыт: пусть имеется ток I и заряд e, неподвижные друг относительно друга. Если они покоятся относительно эфира, то никакого взаимодействия между

ними не будет, но если они движутся, то в первом порядке сила *появится.* В самом деле, при своем движении заряд создает магнитное поле, которое будет действовать на ток. Обратно, из уравнения $\mathrm{div}'\,\mathbf{E}' = \rho - \frac{(\mathbf{J}, \mathbf{w})}{c^2}$ мы видим, что при $\rho = 0$ электрическое поле возникает из-за движения тока и это поле будет действовать на заряд.

Таким образом, сила будет порядка $\frac{Iw}{c}$, т. е. мы в первом порядке получаем отличие от случая покоя, нарушение практического принципа относительности.

Лоренц сразу же разъяснил, в чем здесь дело: при движении тока свободные заряды, перемещение которых по контуру проводника создает I, движутся в магнитном поле и, следовательно, на них, кроме электрического поля, действует и магнитная лоренцова сила. Равновесное распределение зарядов будет поэтому иным: полный заряд по-прежнему будет равен нулю, но в разных местах проводника он будет отличен от нуля. В результате со стороны тока на заряд e будет действовать и магнитное и электрическое поля, причем последнее в первом порядке компенсирует действие магнитного поля. Это кажется случайностью. В каждом опыте приходится отыскивать эту как будто бы случайную компенсацию, и каждый раз она действительно обнаруживается. Это, конечно, неудовлетворительно. В теории относительности все это получается совершенно естественно из общих положений, а здесь — крайне искусственно.

Так или иначе, но Лоренц объясняет все опыты, пока речь идет о первом порядке, т. е. практический принцип относительности в его теории выполняется.

Все было бы хорошо, если бы не новые опыты второго порядка. По Лоренцу, движение Земли во втором порядке должно сказаться на электромагнитных явлениях. Могло случиться, что все останется правильным и во втором порядке. Но оказалось не так. Мы сейчас коротко остановимся на основном опыте второго порядка — опыте Майкельсона.

Луч от источника Q расщепляется полупрозрачной пластинкой D на два луча — $DBDP$ и $DADP$ (рис. 30). Движение Земли через эфир со скоростью w должно сказаться во втором порядке на расположении интерференционных полос в P. Рассчитаем все для пустоты, так как воздух влияет мало.

Для луча DAD, идущего поперек движения Земли, действительный путь будет ct_1, причем $\left(\frac{ct_1}{2}\right)^2 = \left(\frac{wt_1}{2}\right)^2 + l^2$ (рис. 31), откуда

$$t_1 = \frac{2l}{c\sqrt{1-\beta^2}},$$

где $\beta = w/c$. Для луча DBD, идущего вдоль w, имеем время прохождения от D к B: $t' = \dfrac{l}{c-w}$ и от B к D: $t'' = \dfrac{l}{c+w}$.

Таким образом,

$$t_2 = t' + t'' = \frac{2l}{c(1-\beta^2)}.$$

Лучи придут к D с различием по времени во втором порядке

$$t_2 - t_1 = \frac{2l}{c(1-\beta^2)} - \frac{2l}{c\sqrt{1-\beta^2}} \approx \frac{l\beta^2}{c}$$

— совершенно естественный результат, нисколько не противоречащий практическому принципу относительности (замкнутый путь света). Майкельсон в 1881 г. сделал ошибку, полагая $t_1 = \dfrac{2l}{c}$, т. е. считая, что в поперечном направлении луч при ходе туда и обратно перпендикулярен w. Это роли не сыграло, так как это ничего не меняет по существу и лишь оборачивает знак разности хода.

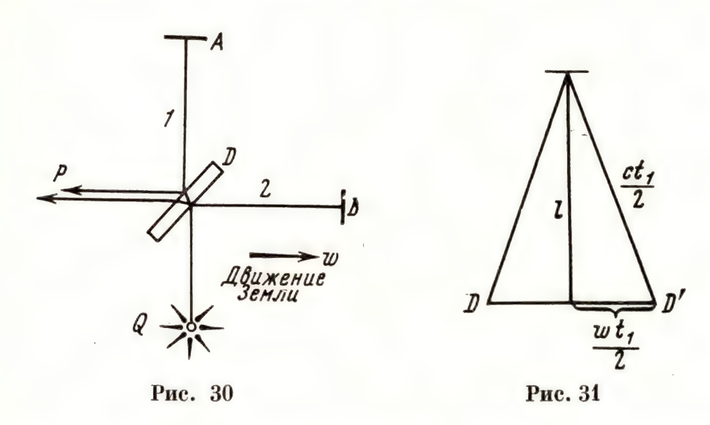


Рис. 30 **Рис. 31**

Если повернуть весь аппарат на 90°, то полосы должны сдвинуться. Опыт крайне прост, в нем нет никакого мудрствования. История этого опыта чрезвычайно интересна и поучительна. В 1878 г. на его идею указал Максвелл, заметив, что если эфир не увлекается, то должен быть эффект во втором порядке, но точность 10^{-8} совершенно недостижима. В 1881 г. Майкельсон подхватил вызов Максвелла, построил в Берлине интерферометр и сделал опыт. Он ожидал смещения в 0,04 полосы, а получил, что смещение, если оно есть, не превосходит 0,015. Это было неожиданно для

самого Майкельсона. В 1887 г. он повторил опыт с большей точностью, и опять — никакого эффекта. Опыт повторялся затем неоднократно. Миллер в 1921 г. получил эффект, но очень малый и никак не связанный с орбитальным движением Земли. Он приписал его движению солнечной системы. Однако условия опыта были ненадежны, и мне кажется, что этому результату нельзя придавать слишком большое значение. С утроенной точностью опыт повторяли Иоос в Германии, Кеннеди в Америке и нашли, что если эффект и есть, то не больше 1 *км/сек* вместо 30 *км/сек*. Теперь можно сказать, что со всей точностью наблюдений эффекта нет; все смещения полос лежат в пределах ошибок опытов.

Вскоре были сделаны и другие опыты второго порядка. К ним относится опыт Трутона и Нобля. Заряженный конденсатор, подвешенный на нити, должен благодаря движению Земли поворачиваться, причем вращающий момент является величиной второго порядка относительно w/c. То, что это так, ясно хотя бы из того, что магнитная сила взаимодействия двух совместно движущихся зарядов пропорциональна w^2/c^2. Каждый из них создает магнитное поле порядка w/c, а в этом поле движется другой, представляющий собой ток того же порядка w/c. Вращающий момент, действующий на конденсатор, получается, согласно теории Лоренца, равным

$$N = \frac{Ww^2}{c^2} \sin 2\varphi,$$

где φ — угол между силовыми линиями и w, а W — электроста-

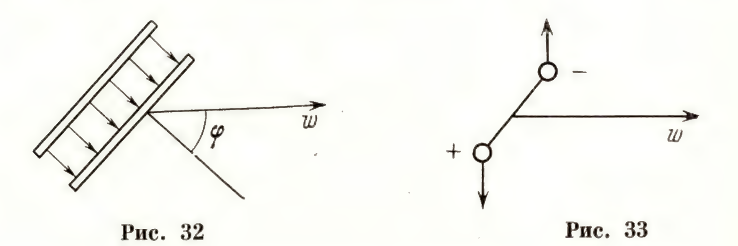


Рис. 32 Рис. 33

тическая энергия конденсатора. Максимальный момент получается при 45° и, как показал расчет, может быть обнаружен экспериментально. Однако эффекта не было. В последнее время с огромной точностью и с тем же отрицательным результатом опыт повторил Томашек. Релей и Брес искали двойное лучепреломление в прозрачных телах, обусловленное движением Земли. С очень высокой точностью опять ничего не было обнаружено.

Итак, при всех доступных точностях не обнаружено никакого влияния движения Земли на электромагнитные и, в частности,

оптические явления. Теория Лоренца была хороша в первом порядке; во втором порядке она отказывается служить.

Уже в 1891 г. Фитцджеральдом, а затем независимо Лоренцом была высказана гипотеза, чуждая всей теории, но дающая объяснение этим отрицательным результатам опытов второго порядка. Эта «гипотеза сокращения» состоит в том, что все тела, движущиеся по отношению к эфиру, укорачиваются в отношении

$$\frac{l'}{l} = \sqrt{1-\beta^2}$$

в направлении движения. Если это принять, то отрицательный результат опыта Майкельсона объясняется сразу: в t_2 вместо l надо подставить l', и тогда

$$t_2 = \frac{2l\sqrt{1-\beta^2}}{c(1-\beta^2)} = t_1,$$

т. е. до второго порядка включительно не должно быть никакого влияния. Принцип относительности при этом *не соблюдается*, ибо хотя разность времен и обращается в нуль, но само время пробега все-таки отличается во втором порядке от времени пробега $\frac{2l}{c}$ в покоящейся системе.

Из создавшегося положения очень трудно было найти выход, но ясно, что гипотеза сокращения придумана уж очень ad hoc. Лоренц пытался, и не без успеха, вложить в нее физическое содержание: размеры тела определяются равновесным расположением составляющих его зарядов. Если в покое мы имеем статическую систему, то при движении относительно эфира появляются магнитные силы взаимодействия зарядов и равновесие будет уже при другом расположении. Если принять лоренцово выражение для силы, то получается как раз то, что нужно — сокращение по направлению движения в отношении $\sqrt{1-\beta^2}$. Однако это не решает вопроса, так как известно, что электростатические силы не могут дать устойчивого равновесия.

В этот период Лоренц, хотя он и физик, выдумал много замечательных математических вещей. Он нашел новую комбинацию, удачно отображающую сокращение и делающую электромагнитные уравнения еще лучше инвариантными. Он предложил преобразование

$$\left.\begin{aligned} x' &= \frac{x - wt}{\sqrt{1-\beta^2}}, \qquad y' = y, \qquad z' = z, \qquad t' = \frac{t - \frac{wx}{c^2}}{\sqrt{1-\beta^2}}, \\ \mathbf{e}' &= \frac{\mathbf{e} + \frac{1}{c}[\mathbf{w}, \mathbf{h}]}{\sqrt{1-\beta^2}}, \qquad \mathbf{h}' = \frac{\mathbf{h} + \frac{1}{c}[\mathbf{e}, \mathbf{w}]}{\sqrt{1-\beta^2}}, \end{aligned}\right\} \quad (11)$$

плюс некоторая формула для скоростей. Здесь уже учтено сокращение. Но опять-таки Лоренц считал временем t и не сделал тех выводов, которые были сделаны потом. Написанное преобразование — это основное преобразование теории относительности. Оно носит имя Лоренца и было написано им в 1904 г. (хотя еще раньше оно было дано Лармором). В вакууме оно дает уже полную инвариантность электромагнитных уравнений. И вот в 1905 г. Эйнштейн все радикально перевернул.

Мы видели, что если преобразование Галилея выражает действительную трансформацию и скорость света не зависит от скорости источника, то надо отказаться от принципа относительности. Обратно: если принять принцип относительности и независимость скорости света от движения источника, то надо отказаться от преобразования Галилея. Эйнштейн знал, что одно с другим совместить нельзя, и он видел, что чем точнее аппаратура, тем достовернее устанавливается справедливость принципа относительности. Независимость скорости света от движения источника тоже не вызывала сомнений. И он сделал вывод, что нужно отказаться от преобразования Галилея. Это не спекуляция. Это плод размышлений над всем ходом развития теории и опытов, это завершение всего предшествующего и выход из той ловушки, в которую попала физика. Другого исхода не было.

Но не является ли отказ от преобразования Галилея таким же злом, как отказ от принципа относительности или от постоянства скорости света? Ведь галилеево преобразование было догмой. Не повлечет ли отказ от него новых противоречий? Громадная заслуга Эйнштейна в том, что он показал, что это не так. Конечно, надо отказаться от старого способа сравнения координат и времени, надо изменить наши представления, но при этом не возникает никаких новых противоречий. Некоторые вещи, о которых мы раньше говорили, как показывает Эйнштейн, не имеют смысла, лишены содержания. Раньше работали расплывчатыми понятиями, укоренившимися навыками, основанными не на логически ясных и определенных утверждениях, а на смутных чувствах. Этому Эйнштейн противопоставляет логически завершенную, до конца ясную и замкнутую систему. В ее создании — его вторая заслуга.

Грубо говоря, это аналогично открытию шарообразности Земли. Когда-то думали, что Земля плоская. Затем пришли к тому, что она круглая, и тогда возник вопрос: как же люди ходят вверх ногами? Это казалось абсурдным, но, подумав, решили, что ничего страшного нет. Увидели, что нет противоречия, нет противоположения каким-либо ясным и четким взглядам, что новые представления противоречат лишь предрассудкам, которые веками оставались незамеченными.

СЕДЬМАЯ ЛЕКЦИЯ

(4.III 1934 г.)

Резюме предыдущих лекций. Постановка вопроса Эйнштейном. Принцип относительности. Вопрос об эфире. Второй постулат Эйнштейна. Противоречивость обоих постулатов. Теория Ритца. Двойные звезды (де-Ситтер) и другие обоснования второго постулата. Из-за чего возникает противоречие между основными постулатами. Вопрос о структуре физических понятий. Определение длины

Резюмируем коротко предыдущее. Мы говорили об электромагнитных и оптических явлениях в движущихся телах. В чем их интерес? Во-первых, эта обширная область явлений, сама по себе интересная для изучения. Во-вторых, здесь, как думали, выявляется взаимодействие материи с эфиром. Существование эфира не вызывало сомнений, и, исследуя оптические и электромагнитные явления в движущихся телах, надеялись получить представление об этой все же загадочной среде. Наконец, в-третьих, мы экспериментируем на Земле, которая движется в мировом пространстве, движется относительно светил. Возникает вопрос: как происходят оптические и электромагнитные явления на движущихся планетах и как они протекают в межпланетном пространстве?

Мало-помалу накопился большой материал, главным образом о явлениях, имеющих место тогда, когда в поле исследования есть предметы, находящиеся в относительном движении (аберрация, допплер-эффект, увлечение Физо — Френеля, магнитное поле движущихся зарядов, ток Рентгена, опыт Вильсона, вся проблема катодных лучей и динамики электрона и т. д.). Параллельно шла и теоретическая обработка. Но еще до подробной теории все больше укреплялось положение, состоящее — без особой точности — в том, что в системе, движущейся по отношению к неподвижным звездам (например, на Земле), оптические и электромагнитные явления протекают совершенно так же, как если бы вся система покоилась. Общее поступательное движение не сказывается на явлениях. Уже Араго пришел к этому заключению, и все дальнейшее его подкрепляло.

Как я уже сказал, наряду с накоплением опыта шла его теоретическая обработка. В сущности, нельзя разделять опыт и теорию: во всяком опыте всегда заключена теория. Но, говоря о теории, я имею в виду попытки построения общей картины, непротиворечиво охватывающей всю совокупность опытных данных. Только при этом условии мы можем удовлетвориться теорией. Вы знаете, что был целый ряд таких попыток, но до 1905 г. они в общем не

увенчались успехом. При создании такой картины всегда возникал прежде всего вопрос о движении эфира, и в зависимости от решения этого вопроса сама задача ставилась по-иному.

Что касается вопроса об электромагнитных явлениях в неподвижных телах, то со времени Максвелла он был однозначно решен. Надо было обобщить теорию на явления, происходящие в движущихся телах. И вот в зависимости от того, считать ли эфир увлекаемым или неувлекаемым, напрашивается то или другое обобщение. Герц, стремясь объяснить независимость электромагнитных явлений от движения системы, попробовал считать эфир полностью увлекаемым. Независимости явлений от движения всей системы он достиг, но в целом попытка не удалась: его теория не объясняла опытов с относительным движением тел, для которых, собственно, она и была создана. Кроме того, принципиальные трудности в вопросе о вакууме и явлениях в очень разреженных телах также вынуждали оставить теорию Герца.

Тогда выступил на сцену Лоренц, который заявил: эфир неподвижен, раз и навсегда. Как же обстоит дело, если, исходя отсюда, строить теорию так, как ее строил Лоренц?

Приняв неподвижный эфир, он отказался от принципа относительности, так как в системе, движущейся через эфир (например, на Земле), явления протекают иначе, чем в системе, связанной с эфиром. Каковы же были результаты его теории? С точностью тогдашних опытов (первый порядок относительно w/c) все объяснялось удивительно хорошо и качественно, и количественно. Утверждение, что общее движение системы не сказывается на явлениях,— это отрицательное утверждение, а, значит, опыт может лишь показать, что с такой-то точностью это верно. Большинство опытов, которые были сделаны в первое время, могли быть произведены лишь с точностью до w/c, да и то ряд мыслимых опытов не был сделан. Но Лоренцу, так сказать в виде бесплатного приложения, удалось показать, что во всех фактически сделанных опытах первого порядка влияния не должно быть и теоретически. Это было как бы «случайное» следствие его теории, вообще не удовлетворявшей принципу относительности. Но «практическому принципу относительности» (понимая под этим фактическую сумму опытов и их фактическую точность) его теория удовлетворяла.

Впрочем, хронологически это не совсем так, ибо уже был опыт второго порядка (опыт Майкельсона), в котором теория Лоренца оказалась несостоятельной: вопреки теории он говорил, что влияния нет. Но теория не должна оказываться несостоятельной ни в одном случае, и поэтому опыт Майкельсона дал теории Лоренца в известном смысле смертельный толчок, укрепляя ту мысль, что принцип относительности больше, чем первое приближение,

и, во всяком случае, выполнен и во втором приближении. Об этом же говорил и опыт Трутона и Нобля. Таким образом, получалось, что принцип относительности действительно есть закон природы и теория Лоренца должна быть оставлена.

Какой же вывод был сделан? Стали думать о том, как видоизменить теорию Лоренца, чтобы охватить и заложенное в ней отрицание принципа относительности, и результаты опытов второго порядка. Вы знаете, что выход указали Лоренц и Фитцджеральд, выдвинув гипотезу сокращения.

Возьмем опыт Майкельсона. Земля движется со скоростью w. Время прохождения света вдоль w есть $t_2 = \frac{2l}{c(1-\beta^2)}$, а время прохождения поперек w есть $t_1 = \frac{2l}{c\sqrt{1-\beta^2}}$. Лоренц предполагает, что длина плеча в направлении w уменьшается: $l' = l\sqrt{1-\beta^2}$. Тогда вместо t_2 будет $t_2' = \frac{2l}{c\sqrt{1-\beta^2}} = t_1$. Это не удовлетворяет принципу относительности, так как не меняется лишь *разность* времен пробега (само же время пробега зависит от движения), однако опыт Майкельсона объясняется.

Конечно, в этой гипотезе много неудовлетворительного, она специально выдумана, но Лоренцу удалось кое-что сделать в направлении ее обоснования.

Что значит, что стержень имеет определенную длину? Это значит, что все частицы находятся в определенных положениях равновесия. Лоренц показал, что при движении относительно эфира электромагнитные силы меняются так, что система в новом равновесии сокращена именно $\sqrt{1-\beta^2}$ раз. Но, во-первых, было известно, что одни лишь электрические силы не дают устойчивого равновесия, а во-вторых, частицы всегда движутся (тепловое движение), так что все это было хорошо, но надлежащего объяснения не давало. Лоренц попробовал органически соединить сокращение со своей теорией и предположил для этого, что электроны сплющиваются при движении и что неэлектромагнитные силы также меняются при движении по тому же закону. Хотя все это и объясняет опыты второго порядка, но слишком все искусственно, принцип же относительности не удовлетворен.

Наряду с этими физическими исследованиями причин сокращения Лоренц сделал математическое открытие: он показал, что уравнения Максвелла (по крайней мере в вакууме), которые при переходе к движущимся осям меняют вид (и тем самым показывают, что независимости от движения нет), остаются инвариантными для некоторого специального преобразования координат, времени и полей, для так называемого лоренцова преобразования. В среде вид уравнений несколько меняется. Этот недостаток ис-

правил Пуанкаре в 1905 г. Он высказал мысль, что надо положить в основу как принцип инвариантность уравнений. Он показал, что лоренцово преобразование, если его несколько иначе истолковать, дает инвариантность и там, где есть материя, но для этого нужно иначе складывать скорости. Надо сказать, что уже в 1900 г. Лармор набрел на лоренцово преобразование; но ни Лармор, ни Лоренц, ни Пуанкаре не считали его физическим преобразованием, а видели в нем только формальный математический прием.

Я уже указал, что лоренцова теория сумела объединить все явления, включая опыты Майкельсона и Трутона — Нобля. Правда, чувство исследователя не было удовлетворено. Принципиальное отрицание принципа относительности шло вразрез с инстинктом, и я лично уверен, что даже вне майкельсоновского опыта почва к признанию принципа относительности была подготовлена. Но как бы там ни было, на майкельсоновском опыте особенно ясно обнаружилась несостоятельность теории неподвижного эфира.

Вот как обстояло дело: громадный материал и невозможность создать общую картину без специально выдуманных гипотез. Правда, в 1905 г. именно благодаря работам Пуанкаре, Лармора, Лоренца и других намечалось проникновение, так сказать, в интимную структуру тел, которое обещало разъяснить положение, но тут появилась знаменитая работа Эйнштейна, которая известным образом поставила все на голову.

Грубо говоря, Эйнштейн сказал: «Вы напрасно хлопочете. В сущности вы хлопочете над очень простыми вещами и сами себе создаете трудности». Нам не стоит вдаваться в то, как далеко зашли Лармор, Лоренц и Пуанкаре, так как постановка вопроса Эйнштейном гораздо яснее и удовлетворительнее. Кроме того, это всегда показатель плодотворности теории — она очень эвристична: на основе постулатов Эйнштейна имеется возможность предсказывать новые явления, и все, что известно сейчас, оправдывает эти предсказания. До Эйнштейна надо было ad hoc придумывать гипотезы. Из теории Эйнштейна, наоборот, вытекают подтверждающиеся на опыте следствия.

Исторически может быть интересно, что работа Лоренца с его преобразованием вышла в 1904 г., затем 5 июня 1905 г. в «Comptes Rendus» появилась работа Пуанкаре, а в сентябре в «Annalen der Physik» — работа Эйнштейна (30 июня была сдана в редакцию). Эйнштейн не знал ни работы Лоренца, ни работы Пуанкаре, но, мне кажется, нет никакого сомнения, что попал в цель всетаки Эйнштейн. Мы приступим теперь к постановке вопроса Эйнштейном, не следуя хронологически, шаг за шагом, но в основных чертах ее развития.

Эйнштейн ставит себе задачей выделить из накопившегося материала бесспорные положения и, исходя из них, посмотреть, к чему они приведут. Эти положения, которые вытекают из всего материала, которые, так сказать, навязаны нам этим материалом, надо положить в основу теории, т. е. надо последовательно их придерживаться и уже на них строить дальше. Такой основой у Лоренца был неподвижный эфир и справедливость в нем максвелл-лоренцовых уравнений. Эйнштейн не старается приспособить эту картину. Вернее, он с ней порывает.

Как первый постулат Эйнштейн выставляет положение: неускоренное движение системы как целого не влияет на законы любых явлений, происходящих в этой системе (ни механических, ни электромагнитных, ни оптических). Для механических явлений это было известно в форме инвариантности уравнений Ньютона при преобразовании Галилея. Трудность была в отсутствии такой инвариантности для электромагнитных уравнений.

Уточним: существует ∞^3 систем отсчета (семейство с тремя параметрами, если отвлечься от поворотов координатных осей и произвола в выборе начала отсчета координат и времени), движущихся друг относительно друга неускоренно, и во всех этих системах явления протекают совершенно одинаково — нет выделенной системы. Не в двух *любых*, а именно *во всем этом классе систем, включающем систему неподвижных звезд*. Мы говорим: есть некоторая система, отнесенная к центру масс неподвижных звезд, и все системы, движущиеся по отношению к ней неускоренно, полностью с ней равноправны. Если же вы возьмете вращающуюся систему и систему, движущуюся по отношению к ней равномерно и прямолинейно, то в них явления не будут протекать одинаково.

Здесь нужно еще одно пояснение. Конечно, явления одни и те же, если условия одни и те же. Если я в одной системе экспериментирую с неподвижной водой, то те же явления получатся в другой системе, если вода неподвижна в ней. При каких же условиях можно назвать системы одинаковыми в отношении состояния материальных тел? Эйнштейн говорит, что для ответа надо учитывать лишь *весомые тела*. А если мы имеем вакуум? Тогда, говорит Эйнштейн, явления всегда одинаковы, вакуум не принимается во внимание, независимо от того, движется система или нет. Если есть две неускоренно движущиеся друг относительно друга системы и в обеих пустота, то явления будут одинаковы. Но ведь имеется эфир? «Я эфира не знаю,— говорит Эйнштейн,— я могу решить, движутся ли *тела* друг относительно друга; движется ли эфир, такого вопроса решить нельзя». Это он выставляет как постулат.

Итак, Герц говорил — эфир увлекается, Лоренц говорил — покоится, Эйнштейн говорит — эфира *нет.* В тождестве двух систем тел, движущихся одна относительно другой, уже заложено то, что мы не можем говорить об эфире.

Когда вообще имеет смысл говорить об эфире? В механических теориях света это имело смысл. Эфир был средой, со всеми свойствами реальных сред. С появлением электромагнитной теории света механическое толкование эфира сделалось тормозом, направляя внимание и усилия на модели, состоящие из шестеренок. Да и принципиально это было неверно, ибо выяснялось все больше, что в основе механических свойств тел лежат их электромагнитные свойства, а тогда не имеет смысла опять сводить электромагнитные явления на механические модели. Но с точки зрения Лоренца или Герца все же можно было говорить об эфире. Он был лишен механических свойств и понимался как носитель электромагнитных полей. Но одно свойство он сохранил: и у Лоренца и у Герца он мог быть *локализован* по отношению к определенной системе. Эйнштейн говорит, что и это свойство отпадает. Можно говорить об эфире, но тогда мы должны примириться с тем, что он покоится во всякой системе (по крайней мере из нашего класса систем), т. е. он лишается последнего свойства, которое позволяло называть его средой,— свойства локализации. Можно называть эфиром пустоту, но это лишено интереса: имеет смысл говорить об эфире лишь постольку, поскольку о нем можно что-нибудь спросить. Но о нем нечего спрашивать. Для физика, стоящего на точке зрения Эйнштейна, понятие эфира теряет смысл.

Но тогда можно, казалось бы, сказать и другое: если эфира нет, то вообще принцип относительности сам собой понятен. Пусть мы находимся в вакууме и эфира нет. В двух системах воспроизведены одни и те же условия, и, поскольку эфира нет, эти системы вообще *ничем* не различаются. Ведь различие раньше было только в движении эфира по отношению к ним. Если хотите, это так. Но *такое* понимание принципа относительности для физика не имело бы смысла. В формулировке же Эйнштейна он отнюдь не сам собой понятен.

Пусть имеется два мира: один весь покоится, другой весь движется со всеми звездами, планетами и т. д. Если бы *так* формулировать принцип относительности, т. е. имеются два мира, которые *полностью* движутся друг относительно друга и в обоих все совершенно одинаково, то это, может быть, и само собой понятно. Но такое утверждение нельзя проверить на опыте, и потому оно неинтересно. Эйнштейн говорит другое: если есть замкнутая система тел и она один раз покоится по отношению к звездам, а другой раз движется, то все будет происходить в ней одинаково, хотя условия не полностью одинаковы, условия различны по от-

ношению к очень удаленным телам. Это уже не самоочевидно. Другими словами, Эйнштейн утверждает, что можно выделить замкнутые (частичные) системы, причем любое неускоренное движение такой частичной системы как целого не влияет на происходящие в ней явления. Если движение ускорено, то это уже не так (например вращение). Никто, однако, не доказал, что ничего не изменится, если вращается как целое вся Вселенная. Когда говорят, что принцип относительности неверен для вращения, то имеют в виду вращение *замкнутой системы*. Вращается ли она или нет по отношению к неподвижным звездам —это вносит разницу, а движется ли она без ускорения или покоится — не вносит. Значит, принцип относительности имеет физический смысл для замкнутой системы. Если бы я постулировал его для мира в целом, то, как сказано, он не имел бы никакого практического, а значит, и физического применения. Он приобретает интерес лишь тогда, когда можно выделить замкнутую систему, для которой он верен. Этот постулат Эйнштейна, который, как он утверждает, осуществляется в природе, не сам собой понятен. Он верен для неускоренного движения, и только. Вопрос о вращении — это вопрос общей теории относительности. Специальная теория относительности рассматривает лишь определенный тип движений — движения неускоренные. Итак, первый постулат Эйнштейна — утверждение правильности принципа относительности в указанном смысле.

Название «принцип относительности» — одно из самых неудачных. Утверждается независимость явлений от неускоренного движения замкнутой системы. То, что это называется «принципом относительности», вводит, как увидим потом, в заблуждение.

Второй постулат Эйнштейна заключается в следующем. Громадный опыт подтверждал правильность концепции Максвелла, которая вообще противоположна концепции классической механики, где есть лишь отдельные конечные тела. Вопрос о *поле* был посторонним в классической механике, понятие поля было чисто математическим. Но в электромагнитных явлениях понятие поля, близкодействие, передача сил от точки к точке — эта концепция основная, и она подтверждалась всем опытом. Точное определение понятия поля заключается в том, что процессы описываются при помощи некоторых величин, удовлетворяющих дифференциальным уравнениям в *частных* производных. Уравнения Максвелла для неподвижной (или для *медленно* движущейся) среды правильны, а из них следует конечная скорость распространения электромагнитных возмущений в пустоте независимо от формы возмущения и независимо от движения источника. Именно в близкодействии и заложена эта независимость. И вот

этот простой факт, подтверждаемый всем материалом, Эйнштейн берет в качестве постулата. Он не знает пока никаких уравнений Максвелла, никакой электродинамики, а только то положение, что скорость света и вообще всякого электромагнитного возмущения не зависит от скорости источника. Это второй постулат Эйнштейна.

Мы все согласны с тем, что оба эти постулата навязаны нам всей картиной, набросанной в прошлых лекциях. Почему же всякому не приходило в голову сказать: давайте их придерживаться и на них строить? Потому, что эти постулаты на первый взгляд (а до Эйнштейна и вообще) казались взаимно противоречивыми.

Естественно поэтому, что, когда Эйнштейн их высказал и они привели к противоречию с тем, что считалось истинным и непреложным, возник вопрос: правильно ли он извлек существенные моменты, которые следует считать постулатами? Может быть, эти положения неверны? Вопрос коснулся именно второго постулата, так как справедливость принципа относительности уже не вызывала сомнений.

И вот тогда Ритц заявил, что он считает второй постулат неверным. Правда, он соответствует духу всей теории поля, но экспериментально он ничем не подтвержден. Он вытекает из уравнений Максвелла, но возможно, что они неверны, и Ритц высказывает гипотезу, что второй постулат неверен. В то время действительно не было прямых экспериментальных доказательств второго постулата, и Ритц был прав, указывая на такую возможность. Мы разберем это и увидим, что в настоящее время имеется достаточно доказательств правильности второго постулата.

Я хотел бы только еще раз подчеркнуть, что Эйнштейн вовсе не стремился придумать теорию относительности в порядке игры ума, как нечто с неба свалившееся. Он действительно стремился выяснить, что является неоспоримым в громадном материале, и ему не оставалось ничего другого, как отсортировать эти два постулата. Поэтому на его теорию нужно смотреть как на единственный исход, который позволил объединить все. И если сразу не назвали эти два принципа, то именно в силу их противоречивости. Чрезвычайная смелость Эйнштейна в том и заключалась, что он этого кажущегося противоречия не испугался.

Почему эти два принципа противоречивы? Достаточно одного примера. Пусть имеем две системы A и B, из которых одна движется прямолинейно и равномерно (пусть B). В некоторый момент наблюдатели A и B находятся друг против друга, и тогда A производит короткую световую вспышку. Через некоторое время сигнал в системе A (неподвижной относительно звезд) будет на поверхности шара радиуса $r = ct$ (рис. 34). Одновременно сигналом будут затронуты все точки на поверхности этого шара. Но в тот

же момент сигнал вышел и для B. Так как справедлив принцип относительности, то наблюдатель B также должен видеть, что через время t сигнал занимает шаровую поверхность радиуса $r = ct$. Но по второму постулату движение источника не должно играть роли. Следовательно, B увидит, что центр сферы находится около него самого, т. е. там, откуда вышел сигнал. Значит, A говорит, что сигнал одновременно захватывает поверхность шара с центром возле него, а B говорит то же, но центр находится возле него, и это должен быть один и тот же шар. Мы приходим к вещи как будто бы невозможной.

Тот же опыт в механике: из точки A во все стороны вылетают осколки гранаты с одинаковой скоростью. Через некоторое время они будут на поверхности шара с центром в A как с точки зрения системы A, так и с точки зрения B. Здесь принцип относительности также верен, но нельзя требовать, чтобы оба наблюдателя видели одну и ту же картину, так как граната покоилась в A, но двигалась в B. Движение гранаты сказывается на скорости полета осколков, т. е. нет независимости скорости движения частиц от источника, значит, нет и противоречия.

Вот почему оба принципа, несмотря на их убедительность и простоту, не были высказаны раньше. Вот почему Ритц ополчился против второго. Ведь достаточно принять, что скорость

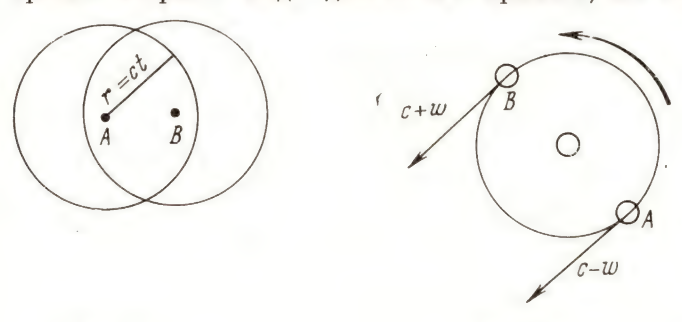


Рис. 34 Рис. 35

света зависит, как и в механике, от скорости источника, чтобы полностью удовлетворить принципу относительности и сохранить преобразование Галилея. Ритц при этом сознательно отвергает уравнение Максвелла.

Существует очень интересная полемика Ритца с Эйнштейном, которая идет, в сущности, глубже. Ритц говорит, что вообще максвелловские уравнения таят в себе известную несообразность: они допускают такие решения, которые, по его мнению, не имеют физического истолкования. Он думает, что эти уравнения дают

слишком много: не только то, что можно наблюдать, но и то, чего наблюдать нельзя. Ритц говорит: непосредственно никто не может сказать, зависит ли скорость света от движения источника, а из совокупности всех опытов этого заключить нельзя. То, что в то время такие опыты уже были, на это никто не обращал внимания. На них в 1913 г. указал де-Ситтер.

Речь идет о двойных звездах. По эффекту Допплера астрономы могут судить об орбитах двойных звезд. Если бы зависимость скорости света от скорости звезды существовала, то это сразу же сказалось бы на наблюдаемых явлениях. Конечно, $w \ll c$, но расстояния огромны и времена прихода света к нам будут очень сильно различаться: звезда послала свет из A, потом пришла в B, и хотя в B она позже, но из B свет идет скорее, и мы могли бы увидеть B и A сразу или даже B раньше A. Во всяком случае, мы увидели бы не так, как оно происходит на орбите. Однако орбиты получаются совершенно правильные, и расчет дает, что если и есть зависимость, то, во всяком случае, меньше 0,002 (т. е. если скорость равна $c + kw$, то $k < 0{,}002$; отрицательное утверждение, конечно, никогда нельзя доказать полностью). Это прямое доказательство было опротестовано Роза, который заявил, что де-Ситтер недостаточно хорошо подсчитал, не учел всех моментов. Теперь можно уверенно сказать, что Роза ошибся.

Можно привести и гораздо более простые соображения, говорящие за постулат Эйнштейна. Возьмем отражение и преломление. На тело падает волна. Почему она отражается? Потому что в теле имеются движущиеся электроны, которые испускают вторичные волны. Если бы скорость этих волн зависела от движения частиц, то они были бы некогерентны между собой и ничего не получилось бы из известных нам явлений отражения и преломления. Конечно, это не доказательство второго постулата, поскольку принимается правильной вся картина преломления и отражения путем вторичного излучения, но все же это аргумент в его пользу [1].

Наконец, опыт Майкельсона объясняется, по Ритцу, тотчас же, если источник движется вместе с прибором, но если в качестве источника взять звезду, то опыт должен был бы дать положительный результат. Томашек в 1926 г. так и сделал, но получил отрицательный результат. В итоге второй постулат нужно признать правильным.

Итак, мы имеем два постулата, на первый взгляд противоречивых, в силу чего их и нельзя было сразу признать.

---

[1] Опыт Физо объясняется в теории Ритца с большой натяжкой. См. *Pauli*, стр. 550.

Но спросим себя теперь: что означает это противоречие? В опыте Майкельсона я наблюдаю принцип относительности, а в опыте де-Ситтера наблюдаю независимость скорости света от движения источника. В том, что я это *наблюдаю*, противоречия нет. Оно наступает в *рассуждениях*. Я определил скорость света, когда я покоюсь. Затем я определил ее, когда я движусь. Если я это действительно сделал и получил одинаковый результат, то в самих фактах противоречия нет. Оно получается тогда, когда говорят: двигаясь, ты не мог определить скорость света как c, двигаясь, ты *должен* был получить $c - w$. Вот в этом рассуждении о том, что *должно* быть измерено, и скрыто противоречие с результатом наблюдений. Не опыты противоречивы, а они становятся такими, потому что вы привносите рассуждение, которому вы абсолютно доверяете. Рассуждение относится к тому, что называется сложением скоростей, оно относится к тому, что если покоящийся человек видел сигнал на поверхности шара, то одновременно другой, движущийся человек *не мог* видеть этот же сигнал на поверхности шара; в том, что *не мог видеть*, а не в том, что видел, заключено противоречие.

Эйнштейн говорит: откуда вы знаете, что преобразование Галилея, т. е. выражаемые им пространственно-временны́е соотношения, правильно? Он выкристаллизовал то звено, где кроется противоречие, это звено оказалось шире, чем все, о чем мы до сих пор говорили. Дело касается не оптики и не механики, а более глубоких вещей — пространственно-временны́х соотношений. Это и придало теории Эйнштейна ее значение. Речь зашла о наших пространственно-временны́х представлениях, о том, что казалось само собой понятным, что казалось «святая святых», чего трогать нельзя, на чем зиждется все естествознание, все наше понимание явлений. И вот сюда Эйнштейн «положил палец».

Нам нужно посмотреть, таким образом, действительно ли правильны те соотношения, которые формулируются галилеевским преобразованием. Для этого надо начать издалека, так как я считаю, что здесь самый гвоздь принципа относительности, что именно здесь большей частью сосредоточено непонимание вопроса, что отрицательно относящиеся к теории относительности не понимают самого подхода.

Итак, наша цель — совместить эти два, казалось бы, противоречивых постулата, показать, что наше рассуждение, приводящее к противоречию, неверно, показать, почему оно неправильно, и заменить его более правильным. Чтобы это сделать, нужно начать с другого принципиального вопроса — вопроса о структуре понятий, с которыми работает физик [1]. Я не могу говорить об

[1] [Говоря о физических понятиях, Л. И. Мандельштам здесь и всюду далее имеет в виду *количественные понятия*, т. е. *физические величины*.]

этом в полном объеме, во-первых, потому, что я не специалист и недостаточно знаю эти вопросы; во-вторых, они нас завели бы слишком далеко в сторону. Но некоторые характерные черты, без которых физик не может работать, нужно усвоить. Мы увидим, что мы говорили массу слов, вообще лишенных содержания, а отсюда и получились все недоразумения. Покажем это на самых простых примерах.

Когда мы говорим, например, о ньютоновых или каких-нибудь других законах, то мы имеем формулы, в которые входят x, y, z. Мы проверяем эти формулы, подставляя вместо x, y, z определенные *числа*. Для этого надо уметь измерять длину.

Что значит для физика измерить длину? Во-первых, надо иметь единицу. Что такое единица длины? Это расстояние между штрихами на стержне, находящемся в Париже, которое называется метром. Можно ли спросить, действительно ли это метр или нет? Нет, нельзя: это по *определению* — метр. В этом сказывается чрезвычайно существенная черта понятий, с которыми оперирует физик. Вообще в логике определить понятие — это значит свести сложное понятие на более простое, т. е. уже известное (не знаю, можно ли это провести до конца). Физик же требует другого: представь реальную вещь. Только так вы свяжете ваше понятие с реальным миром. Нет другого способа определить единицу длины, как показать стержень или другую реальную вещь. Можно сказать: метр — это одна сорокамиллионная земного меридиана. Хорошо, но тогда предмет, который вы предъявляете,— Земля. Вы не можете описать предмет, вы должны его показать; другого способа нет.

Но это не все. Если есть единица, то надо уметь измерять. Я хочу измерить длину стержня. Я укладываю метр. Пусть он уложился 5 раз, тогда я говорю: длина стержня 5 метров. Верно это или нет? Вопрос лишен смысла: это *по определению* так. Если я перенес метр 5 раз, то я говорю, что длина 5 метров. Вы можете сказать, что при переносе стержень мог измениться. Это заявление, взятое само по себе, тоже бессмысленно. Оно имело бы смысл, если бы вы указали, по отношению к чему метр изменился. Тогда я этот другой, «неизменный» стержень возьму за единицу. Мы указали, кроме того, *реальный процесс*, который дает нужное нам число — по определению длину стержня. Этот процесс заключается в перекладывании масштаба. Физик должен иметь «рецепт», как находить длину. Он должен такой рецепт указать, он его не узнает, а определяет. Из непонимания этого вытекают, между прочим, все недоразумения.

В связи с этим мне вспоминается анекдот, который, конечно, никакого прямого отношения сюда не имеет, но, может быть, несколько пояснит мою мысль. Астроном объяснял экскурсии

законы Кеплера, показывал фазы Венеры и т. д. Один из участников, благодаря его, заявил: «Вы очень хорошо объясняли, и я все понял, но вот не понял, откуда вы узнали, что ее зовут Венера?» *Узнать* этого нельзя, можно *назвать*. Я тоже не узнал, что такое длина, а я определил, чтó я назову длиной. Но есть разница: если не Венерой, то называйте как угодно. Определение же длины не может быть вполне произвольным, а должно удовлетворять известным требованиям. Каким?

Смерив ширину шкафа *здесь*, вы находите 1 метр и говорите, что, перенеся шкаф *туда*, вы сможете пронести его через дверь. Откуда вы это знаете? Я убежден, что мой стержень, который я называю метром, если он совпадает с чем-либо *здесь*, то будет совпадать и *там*. Мое определение длины осталось бы правильно и в том случае, если при переносе результат измерения менялся бы. Могло бы быть так, что если я принесу палку прямо в аудиторию, то она совпадет с метром, а если обнесу кругом, то не совпадет. В этом случае обе длины правильны, но «рецепт» *неоднозначен*, непрактичен.

Таким образом, определение равенства двух удаленных отрезков заключает в себе известный рецепт, но оно оправданно, ибо реальные тела, с которыми я оперирую и которые я называю твердыми (например сталь), обладают соответствующими свойствами. Свойства эти таковы, что если два тела совпадают в одном месте, то они совпадут и в другом, и если я одно тело буду сначала обносить кругом, то оно и потом будет совпадать, если совпадало раньше. Я должен иметь такое твердое тело, тело с такими свойствами, чтобы мое определение длины было рационально. Значит, природа не навязывает определений однозначно, но она и не позволяет давать любые определения. Вернее, позволяет, но если я буду определять совершенно произвольно, то я ничего не смогу сделать дальше. Если я хочу строить дальше, то я должен определять, оставаясь в известных рамках.

Я мог бы определить так: если метр укладывается 1 раз, то длина равна 1, если 2 раза, то длина равна 1,5. Запретить так определять нельзя, но это чрезвычайно нецелесообразно: законы Ньютона и вообще все законы были бы совсем другими, было бы много несообразностей и т. д.

Итак, определение основных понятий заключается в том, что я предъявляю определенный предмет, даю определенный процесс и этим предметом и процессом определяю понятие [1], определяю, что я называю длиной и ее измерением. Давая эти определения, я руковожусь требованием однозначности и я выбираю такие

---

[1] [См. примечание на стр. 164.]

тела и процессы, которые ее гарантируют. Гарантируют или нет, *это* уже вопрос опыта.

Вот что физик называет длиной, вот как он находит те числа, которые подставляет потом вместо x, y, z в свои формулы. Если этих рецептов не дать, то формулы станут пустыми, ничего не значащими, физику с ними нечего делать.

Теперь перейдем к времени. Понятие времени t также опирается на определение, базирующееся на предъявлении какого-либо реального процесса. Обычно в качестве часов предъявляется вращение Земли, т. е. считают равными времена, соответствующие равным углам поворота Земли. Это определение. Нельзя спросить: действительно ли Земля в равные времена поворачивается на равные углы? Мы так определили равные времена. Конечно, на деле определение t по вращению Земли делается несколько сложнее, но принципиально положение именно таково: предъявляется Земля и постулируется, что известным углам ее поворота соответствуют такие-то времена. Характер определения остается в силе.

Если такое определение дано, то *узнавать*, вращается ли Земля равномерно или нет,— кажущаяся задача. Этого нельзя сделать, ибо время определено из движения Земли. Что же тогда дают законы Ньютона? Они говорят: если вы определили время, т. е. в любой момент можете найти число t, то вы будете знать, какой функцией от t будут выражаться, например, качания маятника. Т. е. прежде всего нужно определить время, и лишь тогда законы Ньютона приобретают содержание.

Возможен другой способ: считать законы Ньютона справедливыми и определять время из них. Например, в качестве часов предъявляется движение тела по инерции и постулируется, что равным отрезкам пути соответствуют равные времена. Но тогда уже теряет смысл проверка того, является ли движение этого тела равномерным.

Таким образом, в принципе дело не меняется: либо определение времени дается через движение Земли и на опыте находится закон инерции, либо, наоборот, в основу определения кладется закон инерции и с его помощью исследуется движение Земли. Можно, конечно, избрать и какой-либо иной процесс для определения времени t, но, во всяком случае, какой-то реальный процесс указать для этого необходимо.

Непосвященному кажется на первый взгляд, что можно непосредственно узнать гораздо больше, чем это возможно на самом деле. Целый ряд понятий не познается, а *определяется для познания* природы. Эйнштейн показал, что именно этот момент был упущен из виду, и в этом его главная заслуга. Если мы говорим: скорость, одновременность и т. д., то это все пустые слова, пока мы не определили, что они значат. Думали, что в вопросе

об одновременности можно непосредственно сказать, что это такое, а между тем здесь также нужно сначала определить это понятие. До Эйнштейна каждый давал бессознательно свое определение, причем один раз — одно, другой раз — другое. В следующий раз мы и разберем этот вопрос.

ВОСЬМАЯ ЛЕКЦИЯ

(10.III 1934 г.)

Краткое резюме. Вопрос о совпадении разных определений. Одновременность удаленных событий, дискуссия вопроса. Эйнштейновское определение одновременности. Критика возражений против него. Устранение противоречия между основными постулатами, относительность одновременности. О различных определениях одновременности

Резюмирую в двух словах сказанное. На основе всего громадного материала, относящегося к электромагнитным и оптическим явлениям в движущихся телах, Эйнштейн пришел к убеждению, что два положения нужно считать правильными: 1) *принцип относительности*, т. е. признание того, что во всех галилеевых системах координат (галилеевыми системами мы называем класс систем, движущихся друг относительно друга прямолинейно и равномерно) все электромагнитные и оптические явления протекают одинаково. Это именно тот класс систем, в которых, как известно, протекают одинаково механические явления, и в этом смысле Эйнштейн распространил принцип относительности механики на оптические и электромагнитные явления; 2) *независимость скорости света от движения источника*.

Анализируя на основе этих двух постулатов (Эйнштейн выдвигает эти два положения в качестве постулатов) весьма простое явление, а именно распространение мгновенного электромагнитного или светового импульса, выходящего из одной точки, мы пришли к кажущемуся противоречию. При этом анализе мы пользовались пространственно-временны́ми соотношениями, которые дают нам возможность пересчитывать явления из одной системы в другую и которые мы считали само собой понятными. С введением этого звена и получалось противоречие, о котором мы говорили.

Эйнштейн обратил внимание на то, что именно это звено не является само собой понятным, а подлежит более глубокому анализу.

Таким образом, весь вопрос заострился и привел к более глубокому анализу временны́х и пространственных соотношений, т. е. к проблемам, далеко выходящим из области электромагнитных явлений и играющим основную роль в физике. Поэтому и вся теория Эйнштейна далеко вышла за рамки своих первоначальных задач.

Мы начали с того, что с несколько иной стороны посмотрели на самые простые, давно известные соотношения, например на вопрос о длине. Я подчеркнул, что *всякое основное физическое понятие нуждается в определении*. Я не буду повторять все, что было сказано в прошлый раз, и выделю только несколько важнейших моментов.

Во-первых, мы должны иметь *эталон*, должны иметь единицу длины. Для этого нужно предъявить какую-то реальную вещь. Нам нужно создать мост между понятиями, которыми работает математика, и реальными предметами. Для этого нужно показать реальный предмет, например стержень, который находится в Париже и который мы считаем единицей. Затем мы должны дать определение того, что значит *измерить*, т. е. что значит утверждение: стержень, который находится здесь, и стержень, который находится в другом месте, равны. Мы пришли к следующему: если стержень там и стержень здесь совпадают с моим метровым эталоном, причем я один раз прикладываю его здесь, а потом переношу туда и там он снова совпадает, то стержни равны метру. Нельзя спросить, правильно или неправильно мы делаем, поступая таким образом, потому что это есть *определение* длины далеко отстоящего стержня. Точно так же нельзя спросить, действительно ли длина эталона равна метру (конечно, нужно принять во внимание, что температура остается постоянной, но это просто добавочное условие в определении). Я не узнал, что эталонный стержень равен одному метру, а я определил, что́ это значит, и вопрос о правильности и неправильности поэтому неуместен. Это первое: необходимость определения и предъявление какого-то конкретного предмета.

Но наряду с этим не всякое определение допустимо. Если бы я перенес эталон по одному пути и он при этом совпал бы с другим стержнем, а при переносе по другому пути не совпал бы, то наше определение длины все же было бы правомерно. Но оно было бы непригодно. Именно то, что наше определение не зависит от пути, что так обстоит дело в природе для так называемых твердых тел (стали и т. п.), что наше определение оказалось однозначным,— это и сделало его пригодным. Итак, с одной

стороны, определение содержит элемент произвола, с другой стороны, оно плодотворно, потому что дозволено природой в том смысле, что оно удовлетворяет требованиям однозначности, неизменности и т. д.

Я мог бы дать *другое* определение длины. Это несколько сложнее, но я мог бы сказать, что длина волны кадмиевого света равна единице. Это новое определение, но смысл его тот же самый. Я должен был бы показать свет кадмия и сказать: вот это — единица. Я мог бы сделать иначе (и часто это фактически делается), а именно пользоваться другими определениями, которые совпадают с этим. Однако то, что они совпадают,—это уже *показывает опыт*. Одно — первоначальное — *определение* я должен дать. Вы скажете: а астрономические объекты — там ведь вы не прикладываете масштаба? В общем в астрономии все сводится к базе; есть, конечно, и другие предположения, которые усложняют определение, но не меняют дела в принципе.

То же самое и со временем. Я уже говорил, что время мы определяем с помощью какого-нибудь физического процесса, например вращения Земли. Мы можем определять иначе, например при помощи хронометра. Это будет другое определение. Но сказать, что n оборотов Земли и n ударов часов всегда совпадают,— это уже вопрос опыта. Одно из этих положений я кладу в основу как определение; другим можно пользоваться, если я узнал, что это то же самое. Часто пользуются и одним и другим, но это возможно только в результате опытного исследования. Можно в качестве часов взять свет, но опять-таки что-либо одно вы должны определить, а остальное — результат измерения и опыта. Это чрезвычайно существенно. Из дальнейшего вы увидите, почему определение так существенно. Теперь же я укажу лишь на два момента, которые это покажут.

Во-первых, то, что понятие нужно определить,— это просто фактическое положение вещей. Попробуйте истолковать какой-нибудь опыт, попробуйте сказать, что значит какая-нибудь формула, которую вы получаете интегрированием из законов Ньютона. Если вы хотите придать ей какой-либо смысл, какое-либо содержание, то вам неизбежно придется выяснить, что означает x, что означает t, и если вы продумаете вопрос до конца, то вы придете к выводу, что x и t не имеют какого-то свыше данного смысла, а нуждаются в конкретном определении. Вы принуждены дать определения, это не желание, а необходимость.

Второе обстоятельство очень важно, и на нем, по-моему, срывалось все, что было до принципа относительности. Оно заключается в следующем. Если принять, что у нас имеется какое-то априорное понятие, что оно дано „само собой", то легко появляется убеждение, что если, например, я измерю длину двумя способами,

то они *должны* дать одно и то же. Но это неверно. Дадут ли они одно и то же, это *вопрос опыта*, а не долженствования.

Для пояснения я приведу грубый пример. Обычно такие примеры хорошо иллюстрируют лишь одну сторону дела, не охватывая всего, но я надеюсь, что с предлагаемым примером недоразумений не будет.

Когда вы ходите по земле, то кажется совершенно ясным, что значит, что человек ходит вверх головой, и определять этого не нужно. Признаком того, что человек ходит вверх головой, служит то, что его голова дальше от земли, чем ноги. Другой признак тот, что человек, который ходит вниз головой, мне антипараллелен, т. е. если бы я его поставил рядом с собой, то его ноги будут там, где у меня голова, и наоборот. Когда я говорю, что человек ходит вверх головой, то я имею в виду, во-первых, что он мне параллелен, а не антипараллелен и, во-вторых, что его голова дальше от земли, чем ноги. Это совпадает, пока мы все ходим здесь. А антиподы? Почему было возражение, что «там» люди ходят вверх ногами? Потому, что представляли себе сразу и то, что они мне антипараллельны, и то, что их головы ближе к земле. Я даю определение: человек ходит вверх головой тогда, когда его голова дальше от земли, чем его ноги. Будет ли он при этом мне параллелен или антипараллелен,— это вопрос опыта, а не априорного долженствования. Здесь это так, а у антиподов это не так: они мне антипараллельны, но их головы дальше от земли. А ведь единственное, что было бы непонятно и что противоречило бы физиологии человека, это если бы голова была ближе к земле, чем ноги. Но этого никто и не утверждал. Говорили, что Земля не может быть круглой потому, что люди ходят «там» вверх ногами. Этот предрассудок был основан на таком смешении: предполагалось, что «человек ходит вверх головой» есть понятие, свыше данное, а ближе или дальше его голова от земли — это только один из признаков; считалось, что оба признака *должны* идти вместе. Этот предрассудок был основан на непонимании того, что нужно определить, что́ значит ходить вверх ногами, а будет ли это совпадать с другим признаком или нет, — это вопрос опыта. То же самое, как мы увидим потом, произошло с принципом относительности.

То, что я говорил о содержании понятий длины и промежутка времени, знали, конечно, и до Эйнштейна, хотя это не так подчеркивалось, не так вошло в плоть и кровь, как позднее, и не в этих вещах заслуга Эйнштейна. Заслуга Эйнштейна заключается в следующем: мы думали до него, все думали, что основные наши временны́е и пространственные представления исчерпываются этими двумя понятиями, как я их описал, что никаких новых определений давать не нужно. Эйнштейн показал, что это не так.

Для того чтобы действительно иметь достаточно широкие пространственно-временны́е понятия, т. е. такие, которые нам необходимы во всяком вопросе, которыми мы пользуемся во всей физике, этих двух понятий недостаточно. Имеется по крайней мере одно чрезвычайно существенное понятие, которое считалось само собой понятным или вытекающим из указанных ранее, но которое также требует определения, которое без определения не имеет смысла, между тем как им пользовались.

Мы сразу увидим, в чем здесь дело, если рассмотрим простую вещь. Возьмем для простоты определение времени хронометром. Таким образом, время, т. е. то, что я подставляю в формулы Ньютона вместо t, есть то, что показывает стрелка моих часов. Сказать, какое время соответствует происходящему здесь событию, это значит сказать, как здесь стоит стрелка часов, т. е. констатировать совпадение между двумя событиями: положением стрелки на часах и событием, например выстрелом или чем-нибудь в этом роде. Если я говорю, что эти события совпали по времени, то предполагается, что каждый может сразу же сказать, происходят ли эти два события одновременно. Вы знаете, что мы не ошибемся и об одновременности событий, *происходящих в одном и том же месте*, все скажем одно и то же. Поэтому сказать, что нечто происходит в этом месте в 12 часов, это для нас понятная вещь (о констатировании совпадений мы еще поговорим, но пока примем это как данное). Мы можем, следовательно, точно указать, в какой момент отправили сигнал. Мы хотим определить скорость (в механике это самое важное понятие). Пусть движение будет прямолинейным, и на всякий случай предположим, что мы отправляем сигнал из точки A на *большое* расстояние. Мы даем следующее определение скорости: берем какую-то другую точку B, расстояние до которой известно, скажем 1000 *м*, и говорим, что для получения скорости мы должны разделить это расстояние на время. Но как узнать, когда сигнал пришел в точку B?

Я могу принести в точку B хронометр и установить здесь совпадение стрелки с приходом сигнала. Как это сделать? Я возьму часы, такие же, как и в A (как это осуществить, мы не будем обсуждать; мы можем, например, каждые 10 минут выпускать сигнал, и, если часы в точке B также будут показывать, что сигнал приходит через каждые 10 минут, мы скажем, что часы идут одинаково). Но для определения скорости моего сигнала, снаряда, пули этого недостаточно. Я получу, например, что сигнал вышел из A в 12 часов, а пришел в B в 3 часа. Это вовсе не значит, что нужно разделить 1000 на 3. Это было бы так, если бы я был уверен, что в тот момент, когда сигнал вышел из точки A, часы в точке B тоже показывали 12, как и часы в точке A.

Как в действительности я могу это узнать?

Если часы так идут, т. е. если они одновременно в точке A и в точке B показывают 12 и, кроме того, идут одинаково быстро, то я скажу, что они идут *синхронно*. Но как узнать, что двое часов, помещенных в различных местах пространства, идут синхронно, или, что то же самое, как узнать, что два события в различных точках пространства, скажем здесь и на Луне, происходят одновременно? Эйнштейн ответил: вы не можете этого узнать, вы должны дать определение того, что значит, что в двух различных точках часы идут синхронно, что два события произошли в двух таких точках одновременно. Нужно определить, что значит одновременность для различных точек, находящихся на любом (или на достаточно большом) расстоянии, но не совпадающих друг с другом. Другого способа нет.

Об этом никогда не думали, а ведь без знания того, что такое одновременность в двух различных точках, вы не можете сравнивать по времени события, происходящие в различных местах. Между тем этого требует вся механика, без этого не может обойтись физика.

Таким образом, в выяснении того, что понятие одновременности есть такое же понятие, как длина, как время в данном месте, что это третье понятие, которое должно быть определено,— в этом громадная заслуга Эйнштейна. Знали о пространстве, знали о времени, знали в течение многих сотен лет, и никому не приходило в голову, что это так.

Но отсюда вытекают чрезвычайно важные следствия, и мы еще сегодня к ним приступим.

Я хотел бы, однако, подчеркнуть еще раз, что сказанное неопровержимо. Если вы возражаете, то я, в свою очередь, вас спрошу: каким образом можно сказать, не давая определения, идут ли часы синхронно или нет? Это понятие не выводится из других понятий путем каких-либо логических умозаключений.

Все это так. Но отложим пока следствия из этого и обратимся к тем (в известной своей части психологическим) затруднениям, на которые наталкивается указанная постановка вопроса об одновременности.

У меня и у каждого прежде всего возникает следующий вопрос: я говорю, что понятие одновременности требует определения, а с другой стороны, что этого определения не было. Не знали, что нужно определение, значит, не давали его, между тем как без определения нельзя начать, нельзя придать формулам какой-либо смысл. Но ведь нельзя же отрицать, что механика как-то работала и какие-то успехи имела и до Эйнштейна. Ответ прост: давали, в сущности, произвольные определения одновременности, но давали не сознательно. Нехорошо было именно то, что в раз-

личных случаях давались *разные* определения, из чего и возникли неувязки, которые сказались потом.

На некоторых примерах желательно все же проверить, действительно ли это так. Я напомню только, что если с помощью какого-то определения удалось установить синхронность, то вся механика получает смысл. Как определяются длины, мы знаем, мы знаем, как определяется t в данной точке, и если мы установили синхронизм, то мы связали между собой t во всех точках. Принято говорить, что мы таким образом ***хронометризовали*** пространство, т. е. установили, что значит синхронность во всех точках пространства. Я могу тогда сказать, какое время я должен сопоставить событиям в каждой точке пространства, я знаю тогда, что такое скорость, любая зависимость между x, y, z, t приобретает тогда физический смысл.

Я так долго останавливаюсь на этих вопросах именно потому, что в них заключается гвоздь всей теории относительности. Если здесь достичь понимания, то все остальное становится просто само собой понятным и трудности там скорее лишь формально-математические.

Сразу же привлекает внимание, что то противоречие между обоими постулатами, о котором мы говорили выше, связано как раз с одновременностью. Мы ведь говорили, что шар — это место, куда *одновременно* приходит сигнал, пущенный из одной точки. Мы видим теперь, что мы пользовались понятием, которое вообще не было определено, т. е. говорили, в сущности, вещи, которых мы сами не понимали. Поэтому сразу же падает тень на рассуждение, которое привело к противоречию. Очевидно, можно и нужно дать такое определение, чтобы противоречия не было. Это и было сделано Эйнштейном.

Но вернемся к побочным вопросам, которые нам все же кое-что объяснят. Как же поступали раньше?

Я хочу, например, узнать из ньютоновских законов, как падает камень на землю под влиянием постоянной силы тяготения. Я нашел формулу $x = \frac{gt^2}{2}$, но для этого мне нужно было подставлять x и t. Какое t я фактически подставлял? То, которое было на моих часах. Я имел часы только в одном месте и подставлял в качестве t именно те показания часов, которые они давали, когда камень был и здесь и там. Я подставлял показания одних и тех же часов, находившихся в одной и той же точке, а не смотрел, скажем, на часы, которые находились в тех точках, через которые проходил камень. Значит, я подразумевал, что в ньютоновых формулах и во всей механике Ньютона под t нужно понимать показание *часов наблюдателя*. Для малых скоростей и небольших расстояний это не вносит заметной ошибки. Конечно,

вы сразу же скажете, что когда мы смотрим на тело в удаленной точке, то уже прошло некоторое время, которое свет затратил, чтобы оттуда прийти. Таким образом, уже здесь получается неприятность. Фактически я это запаздывание не учитывал, а брал время, которое показывают мои часы (рис. 36). Так поступали всегда.

Когда изучали движение планет, нужно было подставлять в формулы угол и время, и опять подставляли время наблюдателя, время по часам астронома (позвольте мне говорить схематично, не учитывая всех мелочей). Так поступали и поступают и в конечном счете нашли, что... законы Ньютона неправильны. Это нашел Рёмер. Он показал, что если так поступать, то спутник Юпитера не проходит за одно и то же время одно и то же расстояние. Законы Ньютона требовали периодичности затмений, а Рёмер нашел, что ее нет. Он мог это найти, только подставляя какое-то t, и когда он подставлял t, которое давали его часы, он нашел неправильность. Значит, законы Ньютона относятся не к тем часам, которые имел наблюдатель. Тогда Рёмер сказал: законы Ньютона относятся к тем часам, которые находились бы там, на Юпитере. А какие там были бы часы, он посмотреть не мог. Он *постулировал*, что показания часов на Юпитере дают такую разность с его часами, которая, по его мнению, соответствует скорости света. Он нашел, что при его земном времени t законы Ньютона несправедливы, но если подставить другое время, то они будут правильны.

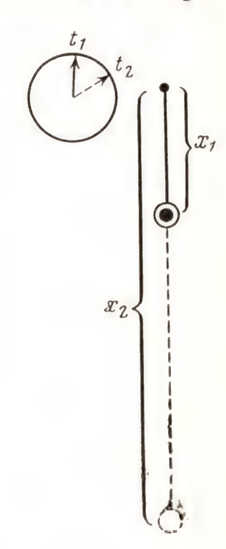


Рис. 36

Я должен указать здесь на некоторую неточность (исторической неточности здесь нет, но хронологическая есть). Рёмер сделал свое открытие в 1675 г., а Ньютон опубликовал свои «Principia» в 1687 г.; поэтому хронологически нельзя сказать, что Рёмер пользовался механикой Ньютона. Но я думаю, что исторически это все-таки будет правильно.

Итак, Рёмер сказал: я буду называть часы синхронными тогда, когда часы в другой точке поставлены на $t_1 = t - \frac{r}{c}$. Если удаленные часы идут так, что я вижу на них время t_1, когда мои часы показывают t, то я буду говорить, что они идут синхронно с моими часами. Это определение. Спрашивается, откуда он знал,

что скорость света до планеты равна c? Он заключил обратное: если взять такое-то значение c, то законы Ньютона будут на Юпитере справедливы. Но тем не менее это было не что иное, как определение. Он *определил*, как нужно поставить часы, чтобы законы Ньютона были справедливы.

Как в действительности определяют одновременность там, где это нужно, например когда пароход уходит в плавание и потребуется находить его долготу? Для этого нужно иметь часы, синхронные с часами на берегу. Как это делалось? Просто на пароходе везли хронометр и считали, что если два хронометра идут совершенно одинаково в одном месте, а затем один из них перевозить, то его показания можно считать синхронными с часами на берегу. Опять это не что иное, как самое грубое определение — определение при помощи перевозки часов. В последнее время стали посылать радиосигналы. Если, скажем, в 12 часов посылается сигнал, то и на пароходе при получении сигнала ставят часы на 12 или учитывают скорость сигнала (конечно, обычно не учитывают). Говорят: будем определять синхронность часов именно так. Это *другое* определение. Наверно здесь было то, что считали, что времена, получаемые при перевозке хронометра и при помощи сигнала,— это одно и то же, что в обоих случаях мы получаем одну и ту же величину. Думали, что если я транспортирую, возможно, хорошие часы, а потом с берега пошлю в 12 часов сигнал, то этот сигнал придет как раз в 12 часов по транспортируемым часам. Именно здесь делали ошибку: это два различных определения — одно при помощи транспортировки, другое при помощи сигнала. Совпадут они или нет — это *вопрос опыта*.

Дорелятивисты думали, что мы узнаем одно и то же время при всяком способе измерения. Точно так же, как в случае с антиподами, считали, что два определения *должны* совпадать. Теория относительности это отвергает. Она утверждает, что если один раз вы перенесете часы, а другой раз будете определять синхронность при помощи сигнала, то получите две *различные* установки часов. Часы, синхронизированные при помощи переноса, окажутся не синхронизированными по радиосигналу. Вы спросите: откуда Эйнштейн это знает? Это другой вопрос. Об этом мы будем говорить потом. Сейчас для нас важно не утверждение Эйнштейна, что мы не получим того же самого, а то, что совпадение или несовпадение — это, во всяком случае, вопрос опыта, а не определения. Вы не можете сказать, что существует какое-то априорное понятие одновременности. Каким-то одним способом вы должны его определить. Но, что покажет другой способ,— это уже вопрос опыта.

Такова позиция теории относительности. Она утверждает, что понятие одновременности является вопросом определения, а

не дано свыше; кроме того, фактически она утверждает, что оба способа дадут различные результаты.

Первого утверждения, мне кажется, никакие будущие опыты не опровергнут. Забыть, что это так, уже нельзя, к прошлому возврата нет. То, что понятие одновременности нуждается, как это указал Эйнштейн, в определении, а не дано свыше,— это шаг, которого взять обратно не сможет никто. Другое, чисто фактическое утверждение теории относительности о том, что два таких определения — одно при помощи транспортировки, другое при помощи сигнала — не совпадают, может оказаться неправильным. Я не исключаю возможности того, что здесь теория относительности может оказаться неправа, хотя сейчас все данные свидетельствуют о правильности этого ее утверждения. Во всяком случае, если теория относительности будет когда-нибудь опровергнута, то это утверждение может оказаться неверным. Но первая часть, которая выяснила структуру понятия, представляет собой логический вопрос и не может быть опровергнута. То, что люди раньше смотрели иначе, а Эйнштейн указал на предрассудок,— этого обратно взять нельзя.

Вы видите, что, несмотря на отсутствие определения понятия одновременности, механика фактически им пользовалась, давая каждый раз свое определение. Долгое время это не вело ни к каким неприятностям. Считалось, что оба указанных ранее определения дают одинаковые результаты. Теория относительности говорит, что они дают различные результаты, но все опыты в механике относились к таким скоростям, при которых эта разница ничтожна и не могла быть замечена. Однако когда подошли к скоростям, близким к скорости света, то не заметить ее уже не могли. Во всей оптике это вопрос уже не последнего десятичного знака, а вопрос существенный.

Вы можете мне сказать: как же работали в оптике? Ведь и в оптике пользовались понятием одновременности, которое также не было определено. На это ответ такой: за исключением рёмеровского опыта, все опыты были таковы, что понятие одновременности в рассуждениях участвовало, а из результата опыта всегда выпадало. Это было понятие, которым в оптике, в сущности, никогда не пользовались, потому что всегда брали путь света туда и обратно. Если пользовались часами, то только часами в одном месте. Правда, для *истолкования* опытов говорили, что свет идет в обоих направлениях с одинаковой скоростью, но в конечном счете это допущение выпадало. Можно спросить: рационально ли, что в обсуждении опытов, из которых понятие одновременности в удаленных точках выпадало, тем не менее им пользовались? Да, это рационально, потому что если устранить эти фактически допускавшиеся рассуждения о движении света в одном направле-

нии с той же скоростью, что и в обратном, то все толкование сделалось бы слишком сложно. Несмотря на то, что оптических опытов с распространением света только в одном направлении нет, в оптике вели рассуждения так, как будто бы они были. Считалось, что отсутствие прямых опытов не принципиальный вопрос, и если в рассуждении пользоваться такими мысленными опытами, то это не такие вещи, которые принципиально не могут быть осуществлены, а такие, которые не осуществлены просто потому, что техника до сих пор до этого не дошла.

В связи с этим я вновь хочу обратить внимание на то, что без понятия одновременности не могло быть понятия скорости, понятие же одновременности должно быть определено. Таким образом, по существу, до Эйнштейна не могли формулировать понятия скорости. Я уже говорил о том, что при обычных скоростях давали другие определения. Но ведь и скорость света была найдена Физо тоже до Эйнштейна. Как же он ее нашел?

Дело в том, что ни Физо, ни Фуко никогда не измеряли скорости света. Они лишь думали, что измеряют ее. Делали так: определяли на одних и тех же часах время между отправлением и возвращением сигнала. При этом не считались с возможностью, что в одном направлении свет может идти быстрее, в другом медленнее. Если *принять*, что время пробега света туда и обратно одинаково и что, следовательно, для определения скорости нужно разделить полное расстояние пополам и разделить на половину промежутка времени между уходом и возвращением сигнала, то тогда они измеряли скорость света. Все это лишь еще

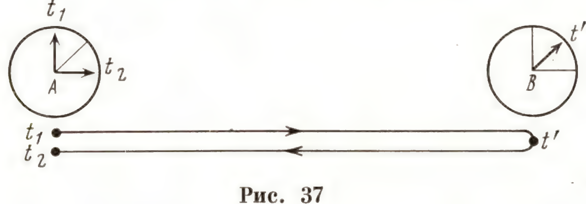


Рис. 37

больше убеждает нас в правильности эйнштейновской постановки вопроса и ни в коем случае ее не опровергает.

Как же Эйнштейн определяет, что̀ называется одновременностью? Эйнштейн дает вполне четкое определение, которое обладает всеми характерными чертами, необходимыми для определения в соответствии с нашими рассуждениями на прошлой лекции. Он указывает реальный физический процесс, который ведет к такому определению. Он говорит следующее. Если имеются часы в одной точке и в другой, то мы посылаем точечный (т. е.

весьма короткий) световой сигнал в вакууме по прямой линии. Пусть часы в точке A при отправлении сигнала показывают время t_1. Пусть время на часах в B, когда сигнал проходит в точку B и отражается в ней обратно, есть t'. Наконец, когда сигнал приходит обратно, пусть часы в точке A показывают t_2. *По определению* часы A и B идут синхронно, если $t' = \frac{t_1 + t_2}{2}$. Повторяю, это определение. Если вы так установите ваши часы, то я скажу тогда, что они идут синхронно. Вы видите, что скорость света и есть тот физический процесс, который служит для определения одновременности.

Таким образом, по Эйнштейну, понятие одновременности тесно связано со скоростью света в вакууме, подобно тому как понятие длины базируется на твердом теле.

Но вы знаете, что определение нельзя выдумать совершенно произвольно, что к нему нужно предъявить известные требования, например требование *однозначности*. Спрашивается, обладает ли это определение таким свойством? Знать это заранее никто не может,— это вопрос опыта. Насколько мы можем судить — прямых опытов здесь нет,— это определение не противоречит требованию однозначности и поэтому дозволено. Однако такое убеждение основано уже на опыте, на физической, реальной природе света.

Итак, однозначность в определении Эйнштейна соблюдена и нужно еще проверить его в отношении других требований. Какие требования мы будем ставить? Прежде всего такое требование: если часы в A и часы в B поставлены синхронно, то не только часы в B должны идти синхронно с часами A, но и часы в A должны быть синхронными по отношению к часам B, т. е. должно быть соблюдено требование *обратимости* синхронности. Если я пошлю сигнал из B в A, то нужно ожидать, что и часы A окажутся установленными синхронно. Легко видеть, что это соблюдено. Именно наша формула $t' = \frac{t_1 + t_2}{2}$ и показывает, что сигнал идет одинаковое время туда и обратно.

Дальнейшее требование заключается в следующем. Если имеются часы в трех точках, A, B и C, и если установлены синхронно часы в A и B и часы в B и C, то и часы в A и C должны быть синхронны, т. е. наше определение должно быть *транзитивно*. Так это или не так, может показать только опыт. Весь имеющийся опытный материал (а мы будем разбирать его с этой точки зрения) говорит за то, что это справедливо, что свет распространяется так, что синхронность часов A и B и часов B и C влечет за собой синхронность часов A и C. Больше того, если синхронность установлена и мы измеряем скорость света (после этого мы уже можем

измерять скорость), то она всегда равна $\frac{r}{t} = c$, какие бы точки мы не взяли. Это и есть постулат постоянства скорости света, который Эйнштейн недаром выделил из всего материала.

Итак, с одной стороны, выбрано некоторое определение синхронности, с другой стороны, оно однозначно, обратимо, транзитивно и имеет место постоянство $\frac{r}{t} = c$. Первое есть утверждение, второе — ссылка на опыт. Может быть, дальнейшие опыты покажут, что последнее неправильно, но не первое. Может оказаться, что требования, которые я предъявляю, фактически не выполняются. Это вопрос природы. Эйнштейн утверждает, что они выполнены и в этом состоит положительное утверждение теории, в этом заключается ее эвристическая сила. Вы должны искать физические способы проверки. Если такая проверка дает отрицательный ответ, то тогда не имеет смысла определять одновременность так, как мы ее определили, потому что это определение не удовлетворяет необходимым требованиям.

Мы увидим далее, что при эйнштейновском определении одновременности, которое согласуется со всем имеющимся у нас опытом, то, что делало противоречивыми оба основных постулата — принцип относительности и независимость скорости света от движения источника,— отпадает. Логических противоречий в них нет. Значит, и в теории относительности логических противоречий нет. На том же опыте, который мы рассматривали раньше, мы сумеем показать, что противоречия здесь нет. Это нетрудно. Но позвольте сделать еще одно отступление.

Понятие одновременности, которое является интегрирующим понятием для времени, слишком важно, чтобы оставить без ответа и другие сомнения, которые высказывались по его поводу. Эти сомнения относились уже не к логической правильности определения Эйнштейна, а к тому, что контроль допустимости, содержащийся в требованиях однозначности, обратимости и транзитивности, недостаточен, что логически такое определение допустимо, но оно настолько несогласно с определенными физическими явлениями, что хотя и допустимо, но не продуктивно.

Те, кто хотел показать, что это определение непродуктивно, исходили не из конкретного физического опыта, а из более общей постановки вопроса. В конечном счете (в той мере, в какой это относится к вопросу о времени) дело сводилось к причинности. Я не могу разбирать вопрос о причинности во всей глубине, а постараюсь изложить его так, чтобы было ясно, о чем идет речь. Говорили так: вся концепция физики требует от понятия одновременности в различных местах, во всяком случае, одного — чтобы закон причинности не был нарушен.

Посмотрим, как это нужно понимать и как отвечает на это Эйнштейн. Представим себе, что мы из некоторой точки A отправляем световой сигнал. Пусть время на часах, находящихся в точке A в момент отправления сигнала будет t_1. Сигнал приходит в момент t' в точку B и затем возвращается в точку A, когда часы в ней показывают t_2. Какое определение дает Эйнштейн? Он говорит, что часы в точке B нужно поставить так, чтобы t' было равно как раз половине $t_1 + t_2$. Это пока что произвольное определение. Я мог бы условиться ставить часы так, чтобы время в B соответствовало какому-нибудь другому моменту в A. И вот, исходя из общих соображений, говорят, что определять одновременность по Эйнштейну нельзя. Почему нельзя? Если я отправляю отсюда какой-нибудь сигнал и он приходит в другое место, в результате чего там что-нибудь происходит, например что-нибудь взрывается, то было бы противно всей концепции физики давать такое определение, которое говорило бы, что при приходе сигнала в удаленную точку время на часах в этой точке окажется более ранним, чем момент посылки сигнала. Если бы мое определение было таким, что я отправил сигнал из A в 12 часов, а он пришел в B по часам в B в 11 ч. 59 м., то такое определение противоречило бы всей физической концепции. Вы можете такое определение дать, но вы ничего с его помощью не сделаете, потому что это значило бы, что я здесь отправил сигнал и в результате моего сигнала взрыв произошел там раньше, чем дан мой сигнал, следствием которого является взрыв. Вся физика построена на понятии

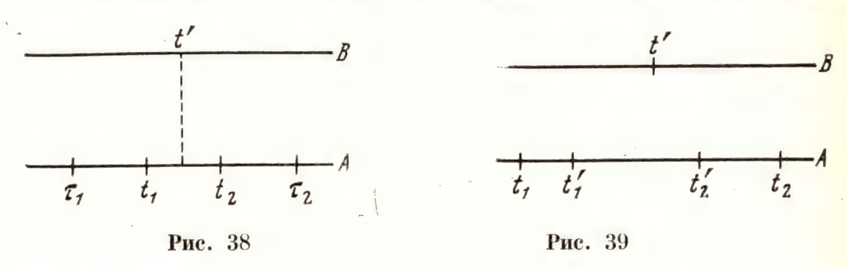


Рис. 38 **Рис.** 39

причинности. Недопустимо строить определение одновременности так, чтобы следствие могло происходить раньше причины.

Это совершенно правильно. Что следует на это ответить? Если бы я дал такое определение, что время прихода сигнала в точку B соответствовало времени τ_1, то это действительно было бы недопустимо, потому что это время более раннее, чем время отправления сигнала. Точно так же нельзя взять τ_2 (отражение сигнала в B после его возврата в A). Но для всех внутренних точек между t_1 и t_2 такое возражение отпадает. При таком опре-

делении всякий сигнал, отразившийся в момент t', соответствующий любой из промежуточных точек между t_1 и t_2, во всяком случае, отразится в B позже, чем вышел из A, и вернется в A позже, чем отразился в B. Значит, чтобы не нарушить принципа причинности, я не имею права поставить часы так, чтобы t' в точке B совпадало с τ_1 или τ_2, но я могу взять любое значение между t_1 и t_2.

Таким образом, требование, чтобы при определении одновременности не был нарушен принцип причинности, может быть удовлетворено неоднозначно. Считая моментом прихода сигнала в точку B любую точку интервала (t_1, t_2), я уже уверен в том, что принцип причинности не будет нарушен. Какую точку я беру, безразлично. Можно взять и среднюю, как это делает Эйнштейн, но ни одна из них не нарушит принципа причинности. Следовательно, то, что необходимость удовлетворить принципу причинности ставит какие-то рамки, правильно, но она не диктует определения однозначно, она оставляет произвол в определении времени на часах в точке B.

Что же было бы, однако, если бы у меня имелся другой способ сигнализирования, более скорый, чем свет? Это значило бы, что такой более скорый сигнал прошел бы туда и обратно в более короткое время, чем световой. Я пришел бы тогда к заключению, что время в точке B можно определить как одновременное с любой точкой более узкого интервала (t_1', t_2'). Наконец, если бы существовал сигнал, который движется с бесконечной скоростью, то интервал между пуском и возвращением сигнала обратился бы в точку и тогда была бы полная однозначность. Значит, если бы существовал сигнал, движущийся с бесконечной скоростью, то требование удовлетворения принципа причинности давало бы однозначное условие и это требование было бы универсальным для всех систем.

Мы видим, таким образом, что обычно высказывавшееся возражение, будто определение Эйнштейна противоречит причинности, неправильно. Оно было бы правильно, если бы существовал сигнал, распространяющийся с бесконечной скоростью. Если бы имелся сигнал, распространяющийся с конечной скоростью, но большей, чем скорость света, то тогда еще была бы неоднозначность, но тогда, если бы я определил одновременность по Эйнштейну, т. е. с помощью света, я нарушил бы либо обратимость, либо транзитивность. Значит, для того чтобы определение удовлетворяло как требованиям обратимости и транзитивности, с одной стороны, так и причинности — с другой, достаточно предположить, что более быстрых сигналов, чем световые, в природе не существует. Если этого не предположить, то эйнштейновское определение не удовлетворит требованиям, которые

мы должны поставить. Либо принцип относительности будет нарушен, либо что-нибудь другое. Если бы существовал сигнал бесконечной скорости, то выполнение принципа причинности влекло бы за собой однозначные условия, мы должны были бы определить одновременность вполне однозначно, и никакого другого определения не могло бы быть. Значит, эйнштейновское определение одновременности предполагает, что *в природе не существует сигналов, которые распространяются быстрее, чем свет.* Это опять физическое утверждение, которое берет на себя принцип относительности [1].

Противоречит ли оно какому-нибудь опыту? Нет. Ни один человек не знает сигнала, не знает процесса, который распространяется с большей скоростью, чем свет. Нет ни одного опыта, который показывал бы, что существует бо́льшая скорость. С другой стороны, уже на этой стадии теория относительности предсказывает, что природа устроена именно так, что такие более быстрые, чем свет, сигналы невозможны, что таких агентов не существует.

Следует уяснить себе, какие скорости, превышающие скорость света, невозможны. Мы вывели это требование из выполнения принципа причинности. Таким образом, нужно уяснить, что не должен быть быстрее света ни один *такой* сигнал, который может *воздействовать.* Если я имею процесс, о скорости которого я могу говорить, но который не обладает тем свойством, что с его помощью можно воздействовать на другую систему, то наличие у такого процесса скорости больше c не противоречит принципу причинности. Принцип причинности нарушается только тогда, когда при помощи данного процесса, имеющего скорость больше c, можно воздействовать, ибо тогда может нарушаться соотношение между причиной и следствием; если же я имею процесс, при помощи которого нельзя воздействовать, то это принципа причинности не нарушает. Такие процессы существуют, но это, как вы видите, не есть возражение. А раньше думали, что это возражение, что это ставит теорию относительности в конфликт с причинностью. О том, какие процессы распространяются быстрее света, мы еще поговорим потом.

Многие стремились дать такое понятие одновременности, которое, как они думали, не зависит от определения, а вытекает из того, что есть какая-то априорная одновременность. Они думали, что они бьют Эйнштейна. Я хочу остановиться здесь на двух возражениях. Одно — серьезное, другое же — элемен-

---

[1] [С дискуссией о роли принципа причинности в определении одновременности и с литературой по этому вопросу можно ознакомиться в книге *А. Гринбаум.* Философские проблемы пространства и времени. М., изд-во «Прогресс», 1969.]

тарное, но и такие возражения делались. Остановимся сначала на втором.

Пусть в точках A и B электрической цепи имеются ключи, а в точке G — гальванометр. Ключи можно замыкать на очень короткое время. Если при замыкании ключей гальванометр покажет отклонение, то это значит, что в обеих точках ключи замкнуты одновременно, ибо в противном случае цепь будет не замкнута либо в одной точке, либо в другой. Отсюда делали вывод, что независимо от каких-либо определений здесь мы констатируем одновременность.

Прежде всего, это не констатация, а определение, а во-вторых, его нет. Вам угодно называть события в A и B одновременными, если гальванометр отклоняется. Это ваше право, вы можете так *определить* одновременность, но не говорите, что вы ее узнаете. Ведь то, что будет идти по проводам, в сущности, будет таким же электромагнитным сигналом (я бы только сказал — испорченным), как и у Эйнштейна. Однако этот электромагнитный сигнал не распространяется с бесконечной скоростью и не дает ничего нового по сравнению с эйнштейновским определением. Элементарная ошибка, в силу которой весь разбираемый «способ» вообще не пригоден, заключается в том, что гальванометр отклонится и при замыкании *одного* ключа, так как волна побежит из-за емкости. Все это возражение возникает из того, что люди представляли себе все устройство имеющим малые размеры и по закону Ома заключали, что, если замкнуть одновременно, будет отклонение, а если не одновременно, то не будет. Но если как следует разобраться, то здесь перед нами тот же электромагнитный сигнал (хотя и по проводам).

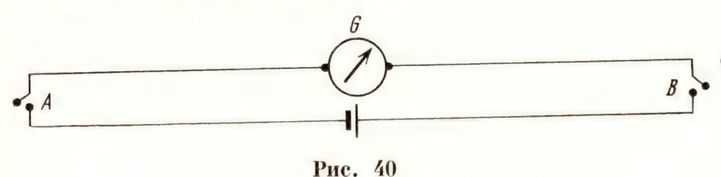


Рис. 40

Другое возражение, которое было сделано, состоит в следующем. Говорили так: возьмите абсолютно твердое тело, например стержень. Пусть в точках A и B находятся часы. Ровно в 12 по часам A я толкну стержень, а так как он абсолютно твердый, то и в другом конце толчок получится одновременно (рис. 41). Следовательно, в точке B я поставлю часы на 12. Таким образом, мы имеем здесь бесконечно быстрый сигнал, с помощью которого можно ставить часы абсолютно синхронно. Но существует ли стержень, который передает толчок с бесконечной скоростью?

Если бы существовал, то такое возражение было бы правильным, мы действительно имели бы сигнал с бесконечной скоростью, и от эйнштейновского определения одновременности пришлось бы отказаться. Таким образом, ответ (его дал Лауэ) прост: если вы толкнете любой реальный стержень, то по нему пойдет акустическая волна, которая идет чрезвычайно медленно по сравнению со скоростью света. Теория относительности утверждает, что таких тел, в которых акустическая волна идет быстрее света, в природе нет и не может быть. В этом теория относительности опять эвристична. Она утверждает, что таких тел не существует. А если таких тел не существует, то это опять-таки не возражение, потому что, как уже было сказано, всякое определение должно относиться к реально возможным телам. Если вы мысленно оперируете с телами, которые противоречат законам природы, то можно сделать что угодно. Это, конечно, не опровержение.

Итак, теория относительности несовместима с абсолютно твердым телом. Она утверждает, что таких тел не может быть. Можно достичь самое большее только того, чтобы скорость звука в теле равнялась скорости света, но не больше, хотя *логически*, конечно, нет ничего недопустимого в том, чтобы скорость звука превысила скорость света. Если вы будете строить что-нибудь (я сейчас стою на классической точке зрения) в расчете на паровую машину, которая дает коэффициент полезного действия, равный единице между температурами 20 и 100°, то вы можете сделать все, что угодно, но природа говорит, что таких машин нет. Точно так же природа говорит, что нет абсолютно твердых тел. Поэтому

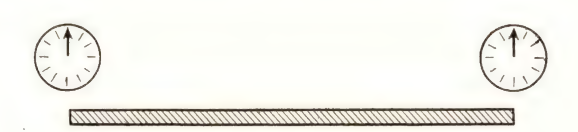


Рис. 41

строить возражение, принимая только принципиально допускаемые, но невозможные в природе тела, нельзя. Если вы спорите, то нужно показать, что такие тела существуют. Теория относительности говорит, что таких тел нет, а вы говорите: я *предполагаю*, что есть, и таким способом хотите ее опровергнуть. Это не возражение. Правда, на этом мы научились новому положению, которого сам Эйнштейн первоначально не заметил, что принцип относительности должен отвергать существование абсолютно твердых тел. В этом достоинство этого возражения.

Еще одно соображение, чтобы вас это не смутило. Все наши определения длины основаны на существовании твердых тел. Но абсолютно твердых тел не существует. Как же теперь быть? Это не существенно. Твердых тел нет, но если перенос совершается достаточно медленно, то тела ведут себя достаточно близко к абсолютно твердым телам. Точно так же жидкость не есть твердое тело, но жидкость может двигаться, как твердое тело. Для определения длины достаточно того, что существуют тела, которые хотя бы при определенных условиях ведут себя, как твердые.

Итак, мы видим, что теория относительности утверждает определенные вещи относительно природы: невозможность скоростей, бóльших скорости света в вакууме, невозможность абсолютно твердых тел. Но в рамках этих утверждений все ее определения логически неопровержимы.

А теперь, в заключение, посмотрим, как действительно обстоит дело с противоречием между принципом относительности и независимостью скорости света от движения источника. Мы увидим теперь, что с точным определением, а не с неопределенными высказываниями здесь никакого противоречия нет. Как мы получали противоречие? Пусть в системе K из точки O в момент t_0 выходит точечный сигнал. Тогда геометрическое место точек, которые одновременно захватываются сигналом в момент t, будет в этой системе шаром радиуса $r = c\,(t - t_0)$. Теперь рассмотрим тот же вопрос в системе K', которая движется по отношению к системе K прямолинейно и равномерно со скоростью v. Пусть сигнал выходит из точки O' системы K' и пусть часы K' в момент выхода сигнала тоже показывают t_0. Принимая во внимание принцип относительности, мы должны сказать, что в момент t геометрическое место точек, одновременно захваченных сигналом в системе K', также должно быть шаром с радиусом, равным $r = c\,(t - t_0)$, но согласно второму постулату с центром в точке O', из которой сигнал исходил в системе K'. Значит, геометрическое место точек, которые одновременно захватываются в системе K, лежит на шаре с центром в O, а точек, которые одновременно захватываются сигналом с системе K',— на шаре с центром в O'. Мы считали, что это противоречие, так как думали, что точки, которые захватываются одновременно в одной системе, и точки, которые захватываются одновременно в другой системе, — это одни и те же точки, а те же самые точки не могут лежать одновременно и на том и на другом шаре. Но теперь мы знаем, что понятие одновременности требует определения и вовсе не означает, что одновременное в системе K одновременно и в системе K'. Что называет Эйнштейн одновременным? В первой системе Эйнштейн называет одновременным достижение сигналом как раз точек, лежащих на шаре с центром в O. Это по определению.

Точно так же во второй системе по определению одновременно захватываются сигналом точки на шаре с центром в O'. Таким образом, по самому эйнштейновскому определению одновременности выходит, что в обеих системах точки, которые одновременно захватываются сигналом, лежат и там и здесь на шаре. Здесь — шар с центром в O, там — шар с центром O'. Но это *не одни и те же точки*, потому что те точки, которые одновременно захватываются в системе K, захватываются не одновременно в системе K', и обратно, потому что понятие одновременности определено в K своим чередом, а в K' — своим чередом, и события, которые являются одновременными в системе K, вовсе не обязаны быть одновременными в системе K'. Эйнштейновское определение одновременности ведет к тому, что два события одновременны не *вообще*, а лишь в какой-нибудь системе. Одновременность является, по теории Эйнштейна, не абсолютной, а относительной. Отсюда ясно, что явление, которое казалось противоречивым лишь потому, что события, одновременные в одной системе считались («само собой разумеется») одновременными и в другой, в действительности никакого противоречия не содержит. Противоречие было обусловлено тем, что рассуждение основывалось на предрассудке. Раз он отпадает, то противоречия нет, и мы видим, что можно совершенно логично построить рассуждение так, чтобы это явление не давало противоречия. События одновременны не абсолютно, а относительно какой-нибудь системы, и тогда все противоречие исчезает.

Итак, понятие одновременности у Эйнштейна удовлетворяет тем требованиям, которые мы предъявляем ко всякому определению. Если пользоваться этим понятием одновременности, то возникавшее ранее противоречие между основными постулатами отпадает, потому что оно было основано на какой-то одновременности, которая считалась само собой разумеющейся. Таким образом, вы видите, что не только в теории относительности нет никаких логических трудностей и темных мест, а наоборот, в прежних понятиях было много непонятного. Те, кто жалуется, что теория относительности что-то переделывает, неправы. То, что утверждает теория относительности, имеет содержание, а то, что говорили до нее, не имело содержания. Понятие одновременности в теории относительности удовлетворяет известным условиям, но это вопрос опыта, в отношении же логики здесь нет никаких трудностей и нет противоречий. Если укажут противоречащий *опыт*, тогда будет о чем говорить, но обычно непонимающие люди заявляют: позвольте, вы перевернули наши понятия о времени и пространстве! Я сказал бы наоборот: у вас не было четких и определенных понятий времени и пространства. Эйнштейн их дал. Может быть, он сделал это так, что эти по-

нятия не описывают природы, но *это* допускает проверку на опыте.

Еще несколько слов. Казалось бы, против теории относительности можно возразить следующее. Хорошо! Если определять одновременность по Эйнштейну, то противоречие между основными постулатами устраняется. Все это очень красиво, но определение Эйнштейна произвольно. Можно определить одновременность и так, чтобы для всех систем она была одна и та же. Я ее так определю и тогда вновь получу противоречие.

Легко видеть, в чем неправильность такого рассуждения.

В эйнштейновском определении есть один чрезвычайно важный момент. Он дает определение в одной системе и совершенно тем же языком дает тождественное определение в другой системе, и тогда *оказывается*, что одновременное в одной системе неодновременно в другой. То, что думали раньше, означает следующее. Представьте себе, что у вас есть одна система и в ней как-нибудь, например по-эйнштейновски, установлена синхронность и вообще вся метрика. Пусть далее имеется другая система. Я мог бы произвольно установить синхронность в этой другой системе таким образом, что часы в ней всегда показывают то же самое время, что и в первой системе. Будет ли такое определение удовлетворять требованиям обратимости, транзитивности и т. п., которые мы выставляем,— другой вопрос (оно не удовлетворяет, но мы примем, что удовлетворяет). Тогда одновременность представляет собой понятие уже для всех одинаковое, т. е. если в этой системе что-либо одновременно, то и в другой оно будет одновременно. Однако теперь световой сигнал не будет одновременно на двух шарах: в одной системе сигнал одновременно будет на шаре, а в другой — не на шаре. Это не противоречит принципу относительности. Почему? Потому что в первой системе мы дали одно определение одновременности, а во второй системе — совершенно другое. Иными словами, во второй системе мы говорим на другом языке, чем в первой. Но тогда нельзя требовать, чтобы был соблюден принцип относительности. Если я требую, чтобы принцип относительности был соблюден, то это вовсе не значит, чтобы он был соблюден на другом языке. Если вы будете называть в одной системе некоторую длину метром, а в другой системе вы будете то же самое называть 4 метрами, то одно и то же явление вы будете описывать различным образом. Когда Эйнштейн говорит, что в природе имеет место принцип относительности, то это значит, что если вы все определения дадите для любых систем одинаково, то явления будут протекать одинаково.

Но еще и не это имеют в виду. В сущности, хотят сказать: меньше разговоров, ближе к эксперименту. У нас есть Земля; сделаем на Земле опыт, скажем тот же самый опыт с распростра-

нением сигнала, и посмотрим, распространяется ли сигнал в виде шаровой волны или нет. Земля движется по своей орбите. Через некоторое время она меняет направление движения. Мы ничего не изменим, аппаратура останется нетронутой, и когда Земля изменит свое движение, мы повторим наблюдение того же самого явления. Что содержится в опыте? Что движение Земли не сказывается, что мы его не замечаем.

Как обстоит дело здесь? Здесь мы даем *третье* определение. Мы имеем здесь две системы: одна система — Земля здесь, и другая система — Земля через полгода. Мы не трогаем наших аппаратов, один раз они на этой Земле, другой раз — на новой Земле, и все, что они показывают, по определению синхронно. Будем ли мы наблюдать при этом то же самое или нет? Вот как ставится вопрос. Теория относительности утверждает — да. Теория неотносительности [1] утверждает — нет. Может быть, права одна, может быть, другая,— это вопрос опыта. Что здесь неправильно и в чем была ошибка? Ошибка дорелятивистов заключается в том, что они неявным образом *определяли* одновременность (абсолютно), но так как они не видели, что одновременность это понятие, которое нужно определить, то они считали, что третье определение — при помощи движения Земли — *должно* давать то же самое. Поэтому они заключали, что в обоих положениях Земли сигнал не может располагаться на шаре. Теория относительности дает свое определение (эйнштейновское) и говорит: то, что вы делаете в опыте на Земле, это третье определение. Что получится,— это вопрос опыта. Я утверждаю, что это третье определение совпадает с эйнштейновским, но, во всяком случае, это дело только опыта. Было бы правильно, если бы дорелятивисты соглашались с тем, что это вопрос только опыта. Тогда одно утверждение выставляется против другого, и опыт должен решить. Неправильность в том, что они говорили: так как существует вполне очевидное понятие одновременности, то всякие его определения *должны совпадать*. Теория относительности говорит: совпадают или нет,— *это вопрос опыта*. Здесь, в сущности, кроется разгадка всех трудностей, именно в убеждении дорелятивистов, что любые определения должны совпадать, от которого они не хотят отказаться. Это тормозит понимание ими теории относительности. Если это уяснить, то можно спокойно отнестись и к одному ут-

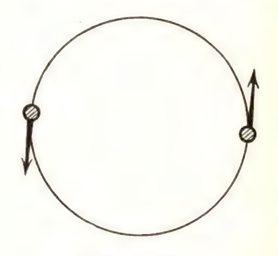


Рис. 42

[1] [Т. е. теория Лоренца.]

верждению и к другому и просмотреть опытный материал. Тогда не будет никаких страстей, никаких недоразумений. То, что дорелятивисты говорят «должно совпадать», есть утверждение, основанное на предрассудке, и это указано Эйнштейном раз навсегда.

ДЕВЯТАЯ ЛЕКЦИЯ

(16.III 1934 г.)

Краткое резюме. Дальнейшая дискуссия понятия одновременности удаленных событий. Вопрос о скоростях, больших c. Установление метрики в инерциальных системах. Вопрос о преобразовании x, y, z и t. Обоснование линейности преобразования. Вывод лоренцова преобразования

Мы видели, что Эйнштейн в своем стремлении найти критерий для объединения всех известных явлений в движущихся телах был принужден обратиться к тому звену рассуждений, которое приводило к противоречию. И он увидел, что здесь оперировали понятиями, которые были недостаточно определены, недостаточно ясны, и в первую очередь — при рассуждениях, которые приводили к противоречию,— пользовались понятием одновременности в различных точках пространства. Эйнштейн обратил внимание на то, что это понятие, как и другие основные физические понятия, нельзя просто принять как нечто данное, что оно подлежит определению. И он дал определение того, что мы будем называть одновременностью для различных точек пространства. Если в точках A и B часы установлены синхронно, то известно, что́ мы называем одновременностью: два события, происходящие в A и B так, что соответственные часы, идущие в этих разных точках, показывают одинаковое время, считаются одновременными. Нужно только знать, как установить часы. Пока часы не установлены, понятие одновременности не имеет никакого смысла. Эйнштейн дает рецепт: пошлите из A световой сигнал в B так, чтобы он скорее всего пришел, проще говоря,— по прямой линии. Сигнал приходит в B и сейчас же отправляется обратно. Пусть моменты выхода и возвращения сигнала по часам A будут t_1 и t_2. Тогда, говорит Эйнштейн, по определению нужно поставить часы в B так, чтобы в момент прихода сигнала в B они показывали $t' = \frac{t_1 + t_2}{2}$. Указав такой способ установки часов, мы опреде-

лили, чтó значит синхронные часы в обеих точках, и тем самым определили, чтó значит одновременность в обеих точках.

Как всякое определение, это определение содержит первоначально элемент произвола. Но оно удовлетворяет всем поставленным требованиям. Во-первых, при таком определении если часы A синхронны с часами B, то и, обратно, часы B синхронны с часами A, что отнюдь не разумеется само собой. Во-вторых, если в трех точках, A, B и C, вы установили часы так, что часы B синхронны с часами A и часы C синхронны с часами A, и если вы захотите теперь таким же способом синхронизировать часы C и B, то вы найдете, что они уже синхронизированы. Другими словами, при этом определении (и при том требовании, чтобы ни одна точка не была исключением) понятие синхронизма есть понятие транзитивное. То, что это так, показывает опыт.

Далее, опыт (по-прежнему в *одной системе*) показывает следующее. Если часы установлены, то приобретает смысл вопрос, чему равняется скорость света; и оказывается, что всегда, в каком бы направлении вы ни производили измерение, вы всегда получите, что между расстоянием от A до B и временем существует зависимость $r = ct$ (конечно, в вакууме) и всегда c будет одна и та же величина — около 300 тыс. *км* в секунду. (При этом мы предполагаем, что гравитация не играет никакой роли. Мы увидим в общей теории относительности, что если делать опыт в гравитационном поле, то этого уже нельзя утверждать.)

Если у меня имеется другая система, которая движется по отношению к первой равномерно и прямолинейно, то в этой системе также следует определить, что значит одновременность, причем мы будем определять по тому же самому рецепту и все будем делать точно так, как в первой системе. Если это сделать, то оказывается, что понятие одновременности не есть абсолютное понятие. Пусть мы установили часы в обеих системах. Мы имеем два каких-нибудь события, происшедших в различных точках. Мы можем установить, одновременны ли они по отношению к первой системе и одновременны ли они по отношению ко второй. И вот при данном выше определении синхронизма оказывается, что если два события в одной системе одновременны, то, вообще говоря, они не одновременны в другой. Одновременность не абсолютна, а относительна.

Если одновременность — относительное понятие, то первое, что мы знаем,— это устранение противоречия между постулатами Эйнштейна. Это основное противоречие обусловливалось тем, что понятие одновременности было абсолютно. Быть может, есть еще какие-нибудь противоречия, но это, во всяком случае, отпадает.

Если спросить себя, из-за чего все-таки в классической физике получалось противоречие, то теперь этот вопрос получает изве-

стный смысл, и вот почему. Как было сказано выше, определение может быть дано так, что одновременность будет абсолютным понятием. Дорелятивистская физика такое определение и давала. Почему же в ней получалось противоречие? Противоречие возникало не из того, что она давала это определение, а из того, что она, давая это определение, навязывала понятию еще и другие свойства, которые в действительности мог указать только опыт. Считалось, что, во-первых, одновременность абсолютна, а во-вторых говорилось, что, *как бы она ни была* определена, она *должна быть* абсолютной. Вот тут была ошибка дорелятивистов. Они, например, считали, что понятие одновременности можно установить переносом часов; они говорили, что если мы возьмем часы и перенесем в другое место, то часы будут идти синхронно, что если мы то же самое сделаем в другой системе, то получится то же самое, и что такое определение будет отвечать абсолютной одновременности. Это неверно. Перенос давал *новое* определение одновременности, и ниоткуда не следовало, что оно совпадает с условиями, которые постулировала дорелятивистская физика. Вот в чем было противоречие, а не в том, что одновременность считалась абсолютной для всех систем.

Если, далее, спросить себя, откуда черпала дорелятивистская механика и вся физика уверенность в том, что одновременность есть такое абсолютное свойство, то, в сущности, корни этого имеются (хотя теперь, может быть, несколько пренебрежительно к этому относятся). Вы видели, что наряду с требованиями, которые мы предъявляем к определению, есть общее требование — принцип причинности, которое запрещает известные определения. Для данного случая можно сказать следующее: нельзя определить одновременность в точках A и B так, чтобы сигнал, пущенный из A, пришел в B раньше, чем он вышел. Значит, часы должны быть синхронизированы так, чтобы сигнал, выпущенный из A, пришел в B в более позднее время. В противном случае, так как при помощи сигнала мы можем влиять на то, что происходит в точке B, мы могли бы влиять действиями, которые происходят позже, на процессы, происходящие раньше. Мы не допускаем этого, и классическая физика тоже, конечно, считала это неприемлемым. Если бы в нашем распоряжении имелись сигналы, обладающие бесконечной скоростью, то тогда уже из требования соблюдения принципа причинности определение одновременности вытекало бы однозначно.

Между прочим, что значит, что сигнал распространяется с бесконечной скоростью, если сигнал нужен только для того, чтобы определить одновременность или синхронизм, а без синхронизма мы вообще скорости определить не можем? Кажется, что здесь логический круг. Но дело в том, что, говоря о беско-

нечной скорости сигнала, я имею в виду, что время, прошедшее от пуска сигнала из A в B до возвращения его в A, должно быть бесконечно малым, должно равняться нулю. Для этого нужно знать только время в одной точке, а не одновременность. Если сигнал имеет конечную скорость, то время t' в точке B нельзя сделать меньше t_1 и больше t_2, оно должно лежать между ними. Если t_1 и t_2 сливаются, то уже нельзя выбирать t' отличным от $t_1 = t_2$. Отрегулировав таким образом часы во всех системах, мы действительно получим абсолютную одновременность. Может быть, до теории относительности эта идея нигде не была так формулирована, но фактически именно так обстояло дело в физике.

Еще одно замечание. Если вторую точку (B) мы будем все более и более приближать к первой (A), то интервал между t_1 и t_2 будет уменьшаться: и если мы подойдем совсем близко, то произвола никакого не останется и будет однозначное определение. Какой бы сигнал мы ни дали, мы должны будем определить одинаково, что значит одновременность для очень близко лежащих точек. Чем ближе лежат две точки, для которых мы определяем одновременность, тем меньше произвол, тем ближе t_1 и t_2, причем они стремятся к совпадению при слиянии точек A и B. Вот почему нет разногласий по вопросу о том, что значит одновременность в двух *совпадающих* точках; мы всегда знаем, что значит, что взрыв совпадает по времени с показанием часов в той же точке. Это мы хорошо знаем, и, какое бы определение мы ни давали, *здесь* это будет одно и то же.

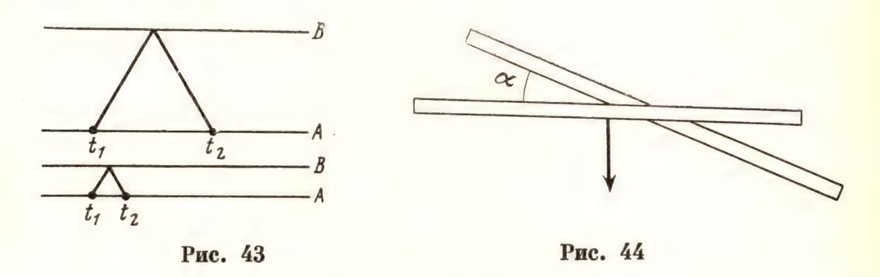


Рис. 43 Рис. 44

Я останавливаюсь так долго на этих вопросах, ибо знаю, что эти вещи наиболее трудны. Я не мог рассмотреть и половины тех вопросов, которые, как показал опыт, обычно возникают; вы сами дадите на них ответы.

Но вот, что я еще хочу сказать. Все наши рассуждения принципиально остаются в полной силе в предположении, что нет скорости, превышающей одну какую-то определенную конечную скорость. Для этой предельной скорости сохраняется все эйнштейновское рассуждение и определение синхронизма, но только

если мы убеждены, что в природе нет скорости, превышающей эту конечную предельную величину. То, что это как раз скорость света — это другой вопрос, к которому мы вернемся потом.

Теперь я хотел бы остановиться еще на одном возражении, которое очень часто упоминается в литературе и о котором очень много говорят. Мы знаем, что существует предельная сигнальная скорость физических агентов. Нам ясно теперь, почему это играет роль и почему наше определение одновременности не противоречит общей картине физики. Потому что нет сигнала, нет воздействия, которые можно было бы передать с большей скоростью, чем скорость света (или вообще некоторая конечная скорость). Однако существуют процессы, для которых формально также можно установить понятие скорости, процессы очень простые, в которых скорость может иметь сколь угодно большую величину. И вот говорят: значит, в природе есть скорости, бо́льшие скорости света, и все ваше построение неправильно. Но вы помните, что нас опровергнуть можно только в том случае, если в природе найдутся процессы *сигнального характера*, более скорые, чем свет. То, что имеют в виду, такого характера не имеет, так что это не возражение (я не могу на этом останавливаться, так как по существу вопрос относится к общей теории относительности, но то, что я сказал, остается правильным).

Возьмем следующий опыт. Представьте себе две линейки, пусть они могут двигаться одна по отношению к другой с известной скоростью (рис. 44). Если угол α достаточно мал, то точка пересечения линеек будет двигаться как угодно быстро. Но это ничего не значит, потому что при заданном движении линеек вы ничего не можете сделать с точкой пересечения. Вы не можете воздействовать на *эту* точку. Другое дело, если вы скажете: первоначально линейка покоится; затем, когда я захочу воздействовать, я начну ее двигать. Точка пересечения начнет двигаться с любой скоростью и будет действительно воздействовать, потому что я привел линейку в движение, когда захотел. Но это как раз и невозможно. Теория относительности утверждает, что этого не может быть. Что произойдет, если вы в некоторый момент начнете двигать линейку? Это тело не жесткое (теория относительности не допускает абсолютно твердых тел), и, пока воздействие дойдет до удаленной точки, пройдет некоторое время. Скорость ни при каких условиях не будет больше, чем c, т. е. вы не сможете за сколь угодно короткий промежуток времени заставить двигаться удаленный конец реального тела. Если кто-нибудь докажет, что в природе это не так, он опровергнет теорию относительности. Однако сделать это невозможно, так как не существует абсолютно жестких тел, т. е. таких, через которые действие распространялось бы с бесконечно большой скоростью. Это ут-

верждение вполне понятно с точки зрения современной молекулярной физики, согласно которой тела состоят из молекул, между которыми действуют электромагнитные силы. Если посмотреть с этой точки зрения, то при деформации вы меняете положения молекул, все ваши действия передаются через электромагнитные возбуждения, и то, что здесь скорость не может быть больше, чем у электромагнитных волн, более или менее понятно.

Наконец, последнее возражение, которое делалось, касается того, что имеет место, например, при аномальной дисперсии, когда коэффициент преломления меньше единицы. Коэффициент преломления есть отношение между скоростью света в вакууме и скоростью в данном теле. Если он меньше единицы, то скорость света в данном теле *больше* скорости в вакууме. Противоречит ли это теории относительности? Когда говорят о коэффициенте преломления, то имеют в виду так называемую *фазовую* скорость. Если имеются установившиеся, синусоидальные волны, то вы можете сказать, что узел волны движется с такой скоростью, которая действительно больше скорости света в вакууме. Но с этим опять ничего нельзя сделать, потому что если имеется такое установившееся движение, то все сказанное справедливо именно и только до тех пор, пока движение установившееся. Как только вы нарушите его, чтобы дать *сигнал*, все построение падает и коэффициент преломления вовсе не показывает, как будут распространяться возмущенные волны.

Иногда говорят, что вопрос решает групповая скорость. Но групповая скорость также может быть больше скорости света. В чем же здесь дело? Дело в том, что и групповая скорость не всегда относится к распространению сигнала. Подробный разбор увел бы нас слишком далеко, но коротко можно сказать одно: нет сигнала со скоростью, большей скорости света. Это отрицательное утверждение, которое трудно поэтому доказать, но в этом сила теории относительности, она делает определенные предположения относительно природы. Пока нет ни одного факта и нет ни одного намека на факт, который противоречил бы утверждению о предельном значении скорости света в вакууме. Всякая теория оперирует с известными гипотезами и ценна постольку, поскольку она не только объясняет то, что известно, но и делает предсказания, так что этого нельзя поставить в минус теории относительности.

После всех этих предварительных рассуждений у нас есть достаточно материала, чтобы построить всю теорию так, как ее строит Эйнштейн. Мы к этому и приступим.

Представим себе, что у нас есть две системы, движущиеся друг относительно друга с определенной скоростью: система K и система K'. В системе K я прежде всего установлю во всех точках

одинаковые часы, что я делаю, пользуясь, скажем, ньютоновыми законами, вообще тем, что мы называем сейчас хронометром. Неважно, будет ли именно такого рода установка, но, во всяком случае, мы воспользуемся тем, что мы называем периодическими процессами; я думаю, что здесь принципиальных затруднений не будет. Итак, мы знаем теперь, как измерить промежуток времени в любом месте. Кроме того, пользуясь твердым телом (скажем, из стали), мы устанавливаем эталон длины. Это может быть парижский метр или какой-либо другой эталон. Мы имеем, далее, рецепт, как измерять любые длины здесь, там и во всяком другом месте. Наконец, по рецепту Эйнштейна мы синхронизируем все наши часы с какими-нибудь одними из них. Теперь все пространство, как говорят, *метризовано.* Каждой точке можно приписать определенные координаты и каждому событию определенное время, вернее, каждому событию можно приписать и определенные координаты и определенное время в нашей системе отсчета. Всякие высказывания, всякие теории, в которые входит зависимость между x, y, z и t, приобретают теперь определенный смысл, мы знаем теперь, что значит t в любом месте, что значит x, y, z.

Точно так же мы все определим в системе K'. По принципу относительности мы должны предположить, что и там дано твердое тело, даны часы и синхронизм установлен посредством тождественного рецепта — при помощи света. Единственное, что будет трудно в первый момент,— это вопрос о том, как взять в K' те же самые эталоны длины и времени. Я беру в системе K определенные эталоны, а какие эталоны я возьму в K'? Пока возьмем любой эталон длины, также сделанный из стали, но произвольный, равно как и любой продолжительности «секунду». Итак, я и в системе K' также метризовал пространство, по тем же самым рецептам. Однако я еще имею свободу в выборе эталонов. Я ограничу ее требованием, чтобы при имеющемся в K' эталоне длины и эталоне промежутка времени скорость света в K' равнялась, как и в системе K, 300 тыс. *км* в секунду. Это требование оставляет произвол в выборе уже только одного эталона, например эталона длины.

Что говорит принцип относительности? Принцип относительности говорит, что если вы облечете в теоретическую форму какое-нибудь явление в одной системе, то соответствующее явление, выраженное в координатах другой системы, будет описываться совершенно так же, т. е. если существует явление, которое описывается в K определенным образом в координатах x, y, z, t, то существует и такое явление, которое совершенно так же описывается в системе K' в координатах x', y', z', t'. «Совершенно так же» — это, конечно, пока трудно утверждать, потому что если

вы выберете эталон в одной системе одним способом, а в другой — несколько другим, то получите и несколько иной результат. Принцип относительности говорит иначе: *наверно, можно* выбрать эталоны так, чтобы было то же самое. При любом выборе эталонов вы не будете иметь того же самого, но принцип относительности говорит: можно выбрать эталоны так, чтобы было то же самое. Это, конечно, совершенно тривиальная вещь [1]. Если я говорю, что в лаборатории здесь и в лаборатории в Ленинграде все явления происходят одинаково, то это возможно только при условии, что я буду одинаково измерять. Если в Москве я буду измерять температуру одним термометром, скажем газовым, а в Ленинграде — ртутным, то я получу различные описания явлений, различные результаты. Что это означает для физики? Это значит, что я получу одни и те же результаты только в том случае, если я буду пользоваться одинаковыми определенными инструментами. Принцип относительности утверждает или предполагает большее — что я *могу* выбрать язык, на котором явления в другой системе описывались бы так же, как и в первой.

Итак, произвольность осталась еще в выборе эталона. Как выбрать одинаковый эталон в обеих системах? Вы увидите, что это просто, и мы вскоре этот вопрос решим.

Какая же теперь перед нами задача? После того как установлена метрика в K и K', перед нами стоит задача преобразовать координаты x, y, z, t в x', y', z', t'. Вот что это значит. Пусть произошло некоторое событие, пусть в системе K это событие имеет координаты x, y, z и случилось в момент t. Это имеет теперь совершенно определенный смысл, и это я наблюдаю. То же самое событие я наблюдаю в системе K', и там оно имеет координаты x', y', z', t'. Спрашивается, как зависят координаты x', y', z', t' от x, y, z, t. Одни координаты являются функциями других. Спрашивается, какими функциями? Мы установили метрику в обеих системах, и, значит, мы можем иметь определенное суждение по этому вопросу. Для того чтобы найти, какова зависимость между теми и другими координатами, мы можем просто прибегнуть к опыту. Это уже вопрос не определения, а измерения. Если я утверждаю, что будут такие-то формулы, то можно это проверить. Можно, например, спросить: правильно галилеево преобразование или неправильно? Пока я не знал, что такое x, y, z, t, этот вопрос не имел смысла. Теперь же вполне можно спросить: дается зависи-

---

[1] [Тривиальным является то, что для выполнения принципа относительности *нужен* одинаковый язык в обеих системах (эта мысль далее и поясняется). Но то, что такой язык *может* быть найден, отнюдь не тривиально. Это ни откуда не следует логически и является вопросом опыта.]

мость между x', y', z', t' и x, y, z, t в форме галилеева преобразования. Правильна она или нет? И можно найти правильную зависимость.

Но весь опыт наш требует — и Эйнштейн это принял,— чтобы зависимость эта удовлетворяла двум постулатам — постулату независимости скорости света от движения источника и постулату относительности. Оказывается, того, что мы знаем, уже достаточно, чтобы получить искомую функциональную зависимость до конца. Таким образом, наша задача будет такова: считая эти два постулата результатами опыта, т. е. принимая их, найти зависимость между двумя четверками величин.

Возьмем какое-нибудь событие, которое мы измерим как в одной системе координат, так и в другой. В качестве события мы выберем приход в определенную точку точечного светового сигнала, выпущенного в момент $t = 0$ из точки $x = y = z = 0$. Это же событие можно рассматривать и в системе K'. Оно будет иметь координаты x', y', z', t'. Если это явление, о котором я только что говорил, то для него в системе K существует определенная зависимость между координатами x, y, z, t, а именно квадрат расстояния от точки O (начала координат), из которой вышел сигнал, равен $x^2 + y^2 + z^2 = c^2t^2$. Такова зависимость между x, y, z и t для этого события, потому что мы взяли распространение светового сигнала, вышедшего из нулевой точки в момент $t = 0$, и принимаем как установленный опытом факт, что свет распространяется с постоянной скоростью c. Но в системе K' это также есть распространение света из определенного источника. Так как по второму постулату скорость источника не играет роли и так как по принципу относительности (первому постулату) я должен видеть в K' совершенно то же самое, что вижу в системе K, то это событие должно дать в системе координат x', y', z', t' ту же самую зависимость, т. е. $x'^2 + y'^2 + z'^2 = c^2t'^2$. Я предполагаю при этом, что в системе K' сигнал выходит из точки $x' = y' = z' = 0$ в момент $t' = 0$. Следовательно, x', y', z', t' должны быть такими функциями от x, y, z, t:

$$\begin{aligned} x' &= f_1(x, y, z, t), \\ y' &= f_2(x, y, z, t), \\ z' &= f_3(x, y, z, t), \\ t' &= f_4(x, y, z, t), \end{aligned}$$

что каждый раз, когда $x^2 + y^2 + z^2 - c^2t^2 = 0$, я получаю $f_1^2 + f_2^2 + f_3^2 - c^2f_4^2 = 0$. Требование, чтобы вторая сумма обращалась в нуль при всех таких x, y, z, t, которые обращают в нуль первую сумму, резко суживает класс возможных функций. Это

требование почти дает однозначность в определении f_1, f_2, f_3, f_4, но не совсем.

В настоящее время известны все классы функций, которые удовлетворяют нашему требованию. Я не буду их перечислять, тем более что не все они для нас приемлемы. Мы наложим еще одно дополнительное требование, которое выделит из общего класса функций, удовлетворяющих нашему условию, более простой класс, которым мы и займемся. Именно, мы потребуем, чтобы эти функции были *линейными* функциями от x, y, z, t.

Чем обосновать нашу уверенность в том, что интересующая нас зависимость выражается линейными функциями от x, y, z, t? Можно дать много обоснований. Этот вопрос, как мне кажется, всегда более или менее правильно излагается во всех учебниках. Я укажу только одно обоснование, которое вполне убедительно и, по-моему, является достаточным. Если мы представим себе в системе K (K — всегда *наша* система, в которой мы производим измерения) какое-нибудь движение, например движение снаряда, и будем считать, что движение равномерное и прямолинейное, то это значит, что x, y, z являются линейными функциями от t. Я спрашиваю затем: как выглядит это движение снаряда, если мы будем рассматривать его в координатах x', y', z', t'? Мне нужно просто подставить функции преобразования, и я получу зависимость между x', y', z' и t', т. е. зависимость в другой системе. И вот естественно требовать, чтобы все прямолинейные и равномерные движения относительно одной системы воспринимались как равномерные и прямолинейные и относительно другой. Следовательно, если я имею линейные уравнения между x, y, z и t, то при подстановке они должны переходить в линейные же, а для этого преобразование должно быть линейным. Правда, оно еще может быть линейным, но не целым, например

$$x' = \frac{\text{линейная функция}}{\text{линейная функция}}.$$

Если мы, кроме того, потребуем, чтобы конечным координатам в одной системе всегда соответствовали конечные координаты в другой, то мы придем к тому, что x', y', z', t' должны быть *целыми* линейными функциями от x, y, z, t.

Здесь есть тонкость, о которой я скажу два слова. Мы требуем, чтобы равномерное и прямолинейное движение в одной системе воспринималось как равномерное и прямолинейное и в другой. Можно ли сказать, что это требование вытекает из принципа относительности? Движение какого-нибудь тела в одной системе, во всяком случае, будет выглядеть иначе, чем в другой, т. е. скорость этого тела в одной системе будет иная, чем в другой.

Мы в свое время говорили, что принцип относительности нисколько не требует, чтобы одно и то же явление всегда выражалось одинаково в разных системах, потому что могут быть различные начальные условия. Я бросаю тело сверху вниз. Для вас это будет вертикальное падение, т. е. прямолинейное движение, а для движущегося наблюдателя это будет парабола. Утверждение принципа относительности касается лишь одинакового протекания явлений при *одинаковых начальных условиях*. Почему же я непременно требую, чтобы прямолинейное и равномерное движение в одной системе было бы таким же и в другой? Будем считать это просто самостоятельным требованием. Часто это выводят из однородности пространства, о которой мы говорили. Это тоже правильная точка зрения, и сводится она к следующему. Вы как-то установили координатные оси одной и другой систем. Если вы одну систему координат перенесете параллельно самой себе, то, конечно, зависимости между координатами станут другие. Требование однородности пространства заключается в том, чтобы никакая точка не была выделена, т. е. если я перенес систему K, то я могу и вторую систему K' перенести таким образом, чтобы между новыми координатами там и здесь установились те же самые зависимости, которые были прежде для неперенесенных систем координат. Это разумное требование, которое говорит, что нет ни одной точки, специально выделенной, что я могу принять за начало координат любую точку и получу то же самое. Если положить в основу это требование, то выходит, что это возможно только тогда, когда функции преобразования линейны, когда

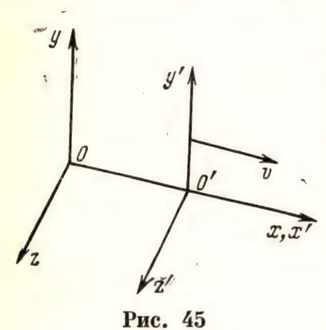


Рис. 45

$$f(x + a) = f(x) + b.$$

Мы потребуем, чтобы это было удовлетворено для всех координат x, y, z, t. Этого, в сущности, достаточно. Все остальное можно просто строго и однозначно вывести. Но, чтобы избегнуть усложнений, мы возьмем обе системы K и K' специальным образом. Это, как вы увидите, никакого ограничения не накладывает, но весь вывод получается нагляднее и проще.

Мы представим себе, что система координат K' взята так, что направление движения системы K' совпадает с осью x. Допустите, что этого нет; тогда можно просто повернуть систему K так, чтобы этого достичь. Так как мы знаем, как переходят координаты из одних в другие при повороте, то пересчитать для этого слу-

чая будет легко. Это мы умеем делать. Поэтому возьмем сразу же систему, в которой движение K' описывается проще и которая отличается от общего случая только тем, что мы определенным образом повернули ее оси. Затем мы примем, что оси y', z' системы K' параллельны соответствующим осям первой системы, а ось x' совпадает с осью x. Кроме того, представим себе, что в момент $t = 0$ начало координат K' совпадает с началом координат K, т. е. при $t = 0$ имеем

$$x = y = z = 0, \qquad x' = y' = z' = 0.$$

Часы в K' мы также можем поставить как угодно. Поставим их так, чтобы часы, находящиеся в начале K', в момент $t = 0$ показывали $t' = 0$. В силу поставленных условий функции преобразования могут быть только однородными, постоянных членов у них быть не может.

Возьмем какую-нибудь точку в плоскости $z = 0$. Так как плоскости xy в обеих системах координат совпадают, то, очевидно, все время, пока $z = 0$, мы будем иметь для этой точки и $z' = 0$. То же самое относится и к координате y. Значит, мы уже наверное знаем, что $z' = \lambda z$. Действительно, z' — линейная функция, а других членов здесь быть не может, потому что при любых x и y как только $z = 0$, то и $z' = 0$. По тем же соображениям $y' = \lambda y$. Ясно, что, ввиду полной равноправности y и z, λ будет одно и то же.

Теперь относительно плоскости $x' = 0$. Какое значение будет иметь x для всех точек плоскости $x' = 0$? Оно будет иметь значение $x = vt$, потому что плоскость $y'z'$ движется параллельно плоскости yz, и когда $x' = 0$, то $x = vt$. Следовательно, $x' = \alpha(x - vt)$.

Наконец, t' будет линейной однородной функцией от x, y, z, t, и я могу ее написать так: $t' = \gamma t - \delta x$. Вы скажете, что я могу прибавить еще $\mu y + \nu z$. Да, можно прибавить, но это должно быть равно нулю. Если бы это было не так, то при $t = 0$ и $x = 0$ мы имели бы разное t' в различных точках плоскости $x' = 0$. Так как все здесь совершенно симметрично, то ясно, что добавок $\mu y + \nu z$ должен равняться нулю. Конечно, все эти соображения симметрии правильны, хотя, возможно, их следовало бы изложить систематичнее и строже. Но все это не имеет большого значения, потому что сделать это очень легко; в литературе это есть, а у нас ушло бы на это много времени.

В конечном счете мы свели дело к тому, что функции преобразования должны иметь вид

$$x' = \alpha\,(x - vt), \;\; y' = \lambda y, \;\; z' = \lambda z, \;\; t' = \gamma t - \delta x,$$

причем все величины $\alpha, \lambda, \gamma, \delta$, вообще говоря, являются функциями

скорости v. Здесь уже вполне ясно, что делала классическая теория. Она принимала в общем случае $t' = t$. В этом сказывалось убеждение, что одновременность имеет одно и то же значение во всех системах.

Таковы общие формулы, к которым мы пришли и которые должны удовлетворять нашим постулатам, т. е. требованию, что коль скоро $x^2 + y^2 + z^2 - c^2t^2 = 0$, то должно быть и $x'^2 + y'^2 + z'^2 - c^2t'^2 = 0$. Первое равенство должно непременно вести за собой второе. Теперь остается сделать немного. Возьмем $x'^2 + y'^2 + z'^2 - c^2t'^2$, подставим наши выражения x', y', z', t' и получим

$$\alpha^2 (x - vt)^2 + \lambda^2 (y^2 + z^2) - c^2 (\gamma t - \delta x)^2.$$

Эта квадратичная функция от x, y, z, t должна равняться нулю, когда $x^2 + y^2 + z^2 - c^2t^2 = 0$. Очень легко показать, что это возможно только тогда, когда эта функция тождественно равняется $x^2 + y^2 + z^2 - c^2t^2$, умноженной на какую-то величину, т. е когда мы имеем

$$\alpha^2 (x - vt)^2 + \lambda^2 (y^2 + z^2) - c^2 (\gamma t - \delta x)^2 \equiv$$
$$\equiv \rho^2 (x^2 + y^2 + z^2 - c^2t^2).$$

Что это значит? В наших уравнениях преобразования величины α, λ, γ, δ не произвольны, а должны быть такими, чтобы выполнялось написанное здесь тождество. А это значит, что должны быть равны коэффициенты при x^2, y^2, z^2, t^2 и xt. Легко видеть, что это даст нам следующие уравнения для α, λ, γ и δ:

$$\alpha^2 - c^2\delta^2 = \rho^2,$$
$$\lambda^2 = \rho^2,$$
$$\alpha^2 v^2 - c^2\gamma^2 = - c^2\rho^2,$$
$$- \alpha^2 v + c^2\gamma\delta = 0.$$

Эти уравнения легко решить. Я просто напишу окончательное решение. Одну из величин можно считать положительной (пока это неважно), скажем ρ мы будем считать положительным. Тогда $\lambda^2 = \rho^2$ (где ρ — еще произвольно), $\alpha = \gamma = \pm \frac{\lambda}{\sqrt{\left(1 - \frac{v}{c}\right)^2}}$, $\delta = \frac{\alpha v}{c^2}$. Таково решение наших уравнений с точностью до знака. Если подставить полученное решение в уравнения преобразования, мы получим окончательный вид этих преобразований. Вместо ρ можно написать λ, это безразлично (позже вы увидите

смысл этого), и пока что просто возьмем положительный знак. Тогда уравнения преобразования получатся в такой форме:

$$x' = \frac{\lambda (x - vt)}{\sqrt{1 - \frac{v^2}{c^2}}}, \qquad y' = \lambda y, \qquad z' = \lambda z,$$

$$t' = \frac{\lambda \left(t - \frac{vx}{c^2}\right)}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Сюда входит еще произвольная величина λ. Это обусловливается тем, что масштабы длины в обеих системах еще не унифицированы. Это и есть лоренцово преобразование с тем отличием, что здесь $\lambda \neq 1$. Мы должны еще оправдать, что $\lambda = 1$. Это мы сделаем в следующий раз.

Мы видим, что если справедливы оба постулата Эйнштейна, то при данном определении одновременности формулы преобразования однозначно определяются. Это лоренцово, а не галилеево преобразование. Почему оно называется лоренцовым? Первым его открыл Лармор, но он ввел его совсем из других соображений. Вы помните, что Лоренц искал, какие нужно взять преобразования для того, чтобы максвелловские уравнения оставались инвариантными. Вы помните, что, исходя отсюда, он нашел такие преобразования, что в вакууме уравнения были инвариантны, а при наличии зарядов — почти инвариантны. Такова была лоренцова точка зрения. Теперь вы видите, что совершенно с другой точки зрения, на первый взгляд никак не связанной с лоренцовой (в действительности связь есть), к этому преобразованию пришел Эйнштейн. Эйнштейн получил его, я бы сказал, без всяких максвелловских уравнений. Если принять два его постулата, то формулы преобразования должны быть именно лоренцовыми.

В следующий раз мы посвятим лекцию дискуссии лоренцова преобразования. Это последний момент, в котором есть многое, о чем следует поговорить и который вызывал много недоразумений.

ДЕСЯТАЯ ЛЕКЦИЯ

(28.III 1934 г.)

Унификация масштабов в двух системах. Вопрос о возможности процесса в вакууме, скорость которого отлична от скорости света и тоже не зависит от движения источника. Длина движущегося масштаба. Критика возражений против релятивистского эффекта сокращения длин. Ход движущихся часов. Часовой парадокс

Мы рассматривали две системы, K и K'. Мы установили в каждой из них понятия координат и времени в любой точке пространства, синхронизировали часы и установили таким образом то, что называется метрикой пространства, причем метрика в каждой системе устанавливается при помощи часов, которые неподвижны по отношению к этой системе. После того как это сделано, все высказывания, в которые входят координаты и время, получают определенный смысл. Наша задача состояла в том, чтобы найти соотношение между координатами и временем в одной системе и координатами и временем в другой. Это понимается так: некоторое событие происходит в одной системе в точке с координатами x, y, z в момент t; спрашивается, каковы координаты x', y', z', t' этого же события в другой системе K', движущейся по отношению к первой равномерно и прямолинейно со скоростью v, т. е. нужно найти, как выражаются x', y', z', t' в функции от x, y, z, t.

Для того чтобы найти формулы преобразования, оказывается достаточным воспользоваться двумя эйнштейновскими постулатами. Насколько мы их уже использовали, об этом мы еще поговорим.

Во всяком случае, мы брали два постулата — постулат относительности и постулат независимости распространения света от движения источника — и пришли к тому математическому требованию, что если мы имеем $x^2 + y^2 + z^2 - c^2t^2 = 0$, то непременно должно быть $x'^2 + y'^2 + z'^2 - c^2t'^2 = 0$.

Формулы преобразования должны быть таковы, чтобы одно было следствием другого, и это сильно суживает класс допустимых функций. Но мы пошли дальше и потребовали, чтобы эти функции были *линейными*. Кроме того, мы ввели специальные координатные системы, ничем не нарушая при этом общности, потому что мы всегда можем повернуть оси в каждой координатной системе. Мы еще вернемся к этому вопросу в другом аспекте. Итак, мы постарались выбрать координатные системы таким образом, чтобы значительно упростить искомые линейные соотно-

шения, а именно так, чтобы эти соотношения, наверное, выражались в форме

$$x' = \alpha(v)(x - vt), \qquad y' = \lambda(v)\, y, \qquad z' = \lambda(v)\, z,$$
$$t' = \gamma(v)\, t - \delta(v)\, x.$$

Нам нужно было найти функции α, λ, γ и δ. Мы нашли их, причем только одна функция осталась неопределенной, и получили

$$x' = \frac{\lambda(v)(x - vt)}{\sqrt{1 - \frac{v^2}{c^2}}}, \qquad y' = \lambda(v)\, y, \qquad z' = \lambda(v) z,$$

$$t' = \frac{\lambda(v)\left(t - \frac{vx}{c^2}\right)}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{12}$$

Такого рода преобразование оказалось единственно возможным.

Когда мы показали, что это единственно возможное преобразование, еще оставались, как вы помните, неопределенными знаки, которые в обеих формулах независимо могли быть и положительными и отрицательными. В формуле для t' мы обязаны выбрать знак «плюс» (считая, что $\lambda > 0$), так как в противном случае направление течения времени («раньше» и «позже») в системе K' будет обернуто. Что касается формулы для x', то мы будем считать, что здесь тоже стоит «плюс». Вот что означает этот выбор. Естественно, что для $v \to 0$ должно получиться тождественное преобразование. Если бы при $v \to 0$ было $x' = -x$, то просто была бы система координат с иначе направленными осями, так что наш выбор никакого ограничения не представляет.

Еще одно замечание. Предположим, что K' движения в направлении отрицательных x. Ясно, что ввиду симметрии это ничего изменить не может, но тогда мы должны написать $\lambda(-v)$. Так как мы признаем, что в пространстве нет никакого выделенного направления, то должно быть

$$\lambda(-v) = \lambda(v).$$

Нет ничего удивительного в том, что у нас остался еще один неопределенный множитель. Наши формулы выражают соотношения между координатами в одной системе и в другой. Но ведь пока мы определили длину в каждой из этих систем, пользуясь масштабами, сделанными в каждой системе особо. И вот теперь, для того чтобы идти дальше, нам нужно перебросить мост между измерениями в одной системе и в другой. Попросту это значит следующее: представьте себе, что у вас есть какой-то масштаб, неподвижный в системе K'. Пусть это будет единичный масштаб.

Я хочу теперь измерить длину этого масштаба, находясь в системе K. По отношению к системе K это будет движущийся масштаб. Значит, задача состоит в том, чтобы измерить длину движущегося масштаба.

Чему же равна длина движущегося масштаба?

Так вопрос ставить нельзя. Вы не можете узнать длину движущегося масштаба, пока вы не определили, что понимается под такого рода длиной.

Под длиной движущегося масштаба рациональнее всего понимать следующее: пусть относительно вас движется произвольно ориентированный масштаб. Вы имеете в своей системе K целый ряд линеек, неподвижных в этой системе, и хотите измерить длину движущегося масштаба. Как это сделать? Когда масштаб проходит мимо ваших линеек, в какой-то момент вы делаете зарубку против его нижнего конца. Но ведь верхний конец тоже движется; значит, вы должны сделать вторую зарубку, и нужно знать, когда ее сделать. И вот самое рациональное определение будет такое: нужно сделать зарубки против концов движущегося масштаба *одновременно* по отношению к системе, в которой вы измеряете, скажем сделать зарубки в 12 часов, т. е. когда часы системы K показывают в обеих точках 12. Таким образом, измерение длины движущегося масштаба всецело будет зависеть от того, как синхронизированы часы в системе K.

Здесь возникает обычно много недоразумений. Нет понятия длины движущегося масштаба самой по себе. Этому понятию мы даем определение такое, что в один и тот же момент по часам измеряющей системы нужно от обоих концов движущегося масштаба сделать зарубки на неподвижном масштабе. Это не что иное, как обобщение того, что мы говорим обычно, когда масштаб неподвижен в нашей системе. В этом частном случае мы также говорим: чтобы измерить неподвижный масштаб, я должен сделать зарубки на линейке на обоих концах. Но здесь безразлично, сделаны ли они одновременно или разновременно, потому мы об этом и не говорим. Мы можем сделать зарубки одновременно, но можем сделать вторую и позже. В случае же движущегося масштаба от разновременности получатся другие результаты. Поэтому необходимо было дать новое определение. Оно не вытекает из обычного, но ему не противоречит и содержит его как частный случай.

Измерим теперь из неподвижной системы K единичный эталон, покоящийся в K' и расположенный перпендикулярно к направлению движения, скажем вдоль оси y. Что значит, что эталон K' равен единице? Это значит, что $y'_2 - y'_1 = 1$. Я должен на моем неподвижном масштабе сделать зарубки в один и тот же момент. Но как раз для перпендикулярного направления время

t в мои преобразования не входит, и поэтому в данном случае безразлично, сделаны ли зарубки одновременно или нет. Что же я буду измерять в моей системе K? Здесь я буду называть длиной масштаба расстояние $y_2 - y_1$. По формулам преобразования (12) можно написать, что

$$y_2 - y_1 = \frac{1}{\lambda(v)}(y_2' - y_1').$$

Значит, эталон, который я условился считать единицей в K', в системе K будет иметь длину $l = 1/\lambda(v)$.

Теперь проделаем такое же измерение, но из системы K' эталона, который в K выбран за единицу. Этот эталон, измеренный из K', согласно той же формуле преобразования $y' = \lambda y$, будет, очевидно, иметь длину $l' = \lambda(v)$. Значит, при произвольно взятых масштабах мы скажем, что эталон, который в K' считается единицей, будучи измерен из K, оказался равен $1/\lambda(v)$. Это и не удивительно, поскольку масштабы были разные, поскольку там единицей является один сантиметр, а здесь, скажем, два.

Но плохо не это. Плохо то, что если взять единицу K', то в K она будет, скажем, уменьшена, а единица системы K в системе K' будет увеличена. Значит, говоря на таком языке, мы не получаем принципа относительности, потому что эталон одной системы, измеренный по отношению к другой, укорочен, а эталон другой системы, измеренный по отношению к первой, удлинен. Однако принцип относительности и не говорит, что при *любом* языке все будет одинаково. Он говорит только, что всегда *можно* подобрать соответствующий язык. Следовательно, мы должны подобрать такой язык, чтобы эталон, измеренный из одной и из другой системы, был одинаков. Отсюда требование $\frac{1}{\lambda(v)} = \lambda(v)$, т. е. $\lambda^2(v) = 1$. Значит, последнее, чем мы могли распоряжаться — произвольный выбор эталона, мы сделали и сделали так, чтобы принцип относительности был соблюден. Тогда формулы преобразования становятся уже вполне однозначны и принимают вид

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}, \qquad y' = y, \qquad z' = z, \qquad t' = \frac{t - \frac{vx}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{13}$$

Что можно сказать относительно этих формул? Мы брали независимые эталоны в K и K', а потом так их подогнали, чтобы принцип относительности был удовлетворен. Принцип относительности говорит, таким образом, что если вы возьмете в системе K два совершенно одинаковых эталона и один из них передадите

в K', то в K' этот эталон совпадет с той единицей, которую мы описанным путем там установили. Значит, эталон во второй системе я мог получить тем, что эталон, который в K был равен другому такому же экземпляру, я просто передаю. Это не следует прямо из преобразования (12). При его выводе все было сделано независимо в каждой системе относительно этой системы; теперь же я могу говорить о передаче, о транспорте. То же самое и с часами. Часы, будучи переданы, также дадут то время, которое было нами определено [1].

Отмечу еще раз, что Эйнштейн приходит к преобразованию (13), совершенно не касаясь максвелловских уравнений, исходя только из своих двух общих постулатов, в которых, правда, скорость света играет очень существенную роль. Это, конечно, не случайно, и мы увидим, какой глубокий смысл это имеет с точки зрения Эйнштейна.

Я хотел бы обратить внимание еще на одно обстоятельство. Мы искали преобразование так, чтобы из $x^2+y^2+z^2-c^2t^2=0$ следовало, что и $x'^2+y'^2+z'^2-c^2t'^2=0$. Но мы вовсе не требовали, чтобы в том случае, когда $x^2+y^2+z^2-c^2t^2=a$, при подстановке вместо x, y, z, t координат x', y', z', t' это выражение сохраняло то же самое значение. Мы полагали

$$x'^2+y'^2+z'^2-c^2t'^2=\rho^2(x^2+y^2+z^2-c^2t^2),$$

но оказалось, что

$$\rho^2=\lambda^2=1.$$

Таким образом, если $x^2+y^2+z^2-c^2t^2=a$, то при подстановке x', y', z , t' вместо x, y, z, t мы получим опять $x'^2+y'^2+z^2-c^2t'^2=a$, т. е. эта квадратичная форма инвариантна по отношению к лоренцовому преобразованию. Это больше, чем мы требовали.

Из формул преобразования мы видим далее, что, действительно, во всей теории Эйнштейна скорость света в вакууме (в отсутствие гравитации) есть предельная скорость. Скорость материальной системы не может быть равна или больше c. Если бы это было так, то под знаком корня получилась бы отрицательная величина и x' и t' были бы мнимыми, т. е. никакого физического содержания мы этим формулам не придали бы. Мы утверждаем, что такие скорости невозможны.

Интересен следующий вопрос. Может ли существовать какой-нибудь другой процесс в вакууме, который распространялся бы

[1] [Передача эталонов из одной системы в другую связана с ускорением, которое должен претерпеть эталон. В пределах специальной теории относительности мы ничего не можем сказать о влиянии этого ускорения. Предыдущее рассуждение, которое дало нам $\lambda=1$, показывает, что сверить эталоны можно и без передачи, так сказать «на ходу».]

независимо от движения источника, как и свет, но с другой скоростью. Если теория относительности правильна, то это невозможно, потому что если бы это имело место, то мы путем совершенно таких же рассуждений пришли бы к тому, что для выполнения принципа относительности и принципа постоянства скорости нашего агента формулы преобразования должны иметь тот же самый вид, но только вместо c в них должна входить скорость этого нового агента. Свет существует — мы это знаем. Значит, для того чтобы удовлетворить принципу относительности по отношению к свету, формулы должны быть такими, как они написаны, а для того чтобы удовлетворить ему для другого агента, они должны быть такими же, но с другим значением c. Значит, если бы существовал другой агент, то для него принцип относительности был бы нарушен. Для $v = 0$ формулы «преобразования» совпадали бы, но для других скоростей, если принцип относительности верен для c, для другого агента он был бы неверен. Таким образом, принцип относительности исключает возможность существования других агентов в вакууме, которые распространялись бы с другой скоростью, чем c, не зависящей от движения источника.

Далее, из формул (13) мы видим то, что мы уже вложили туда, а именно относительность одновременности. Пусть что-нибудь происходит в точках x_1 и x_2 в системе K и пусть эти два одновременно происходящих события соответствуют времени t. Каковы соответственные t', т. е. каковы показания часов для этих событий в K'? Это легко определить. Нужно подставить t и x_1 для того, чтобы узнать t_1' для первого события, и t и x_2, чтобы найти t_2 для второго. Мы видим, что t_1' и t_2' отличны друг от друга, т. е. мы убеждаемся в том, что если в системе K в двух различных точках одновременно произошли два события, то с точки зрения системы K' они произошли не одновременно. Одновременность есть понятие относительное. Если два события одновременны в одной системе, они, вообще говоря, не одновременны в другой.

Разумеется, с точки зрения принципа относительности можно принять за исходную систему не K, а K'. Поскольку мы уже знаем лоренцово преобразование, мы можем переписать его так, что вместо x будет стоять x' и обратно, но только скорость будет тогда равна $-v$, если раньше она была v. Таким образом, если мы исходим из системы K', то мы напишем лоренцово преобразование так:

$$x = \frac{x' + vt'}{\sqrt{1-\frac{v^2}{c^2}}}, \qquad y = y', \qquad z = z', \qquad t = \frac{t' + \frac{vx'}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}. \tag{14}$$

С другой стороны, мы могли бы просто вычислить x из линейных уравнений (13), и это должно было бы дать то же самое. Нетрудно убедиться, что так оно и получается, т. е. формулы преобразования такого рода взаимности удовлетворяют.

Перейдем теперь к одному из важнейших следствий, которое вытекает из лоренцова преобразования и которое является, как утверждают, одним из самых больных вопросов, хотя, с моей точки зрения и с точки зрения всякого, кто этим занимался, оно не представляет никаких затруднений.

Речь идет об измерении длины движущегося масштаба, но в том случае, когда этот масштаб расположен параллельно направлению движения. По-прежнему имеются системы K и K'. В K' покоится масштаб, длина которого равна $l_{K'}^{K'} = x_2' - x_1'$ (нижний индекс при l будет показывать, в какой системе масштаб покоится, а верхний — в какой системе производится измерение). Спрашивается, какова будет длина масштаба, если измерить ее из системы K? Мы знаем теперь, что это значит. Нужно найти соответствующие координаты x_1 и x_2 концевых точек в одно и то же время, т. е. при условии, что часы системы K в обеих точках показывают одно и то же, а потом измерить расстояние между x_1 и x_2. Это по определению есть то, что мы называем длиной движущегося масштаба: $l_{K'}^{K} = x_2 - x_1$. Каковы же будут координаты этих двух точек в системе K в один и тот же момент? Мы должны написать формулы преобразования

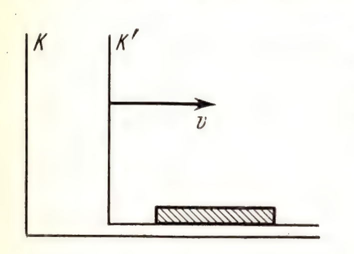


Рис. 46

$$x_{1,2}' = \frac{x_{1,2} - vt}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Если произвести вычитание, то, так как по условию t в обоих случаях одно и то же (по определению длины движущегося масштаба мы измеряем для одного и того же показания часов t), получится

$$x_2 - x_1 = (x_2' - x_1') \sqrt{1 - \frac{v^2}{c^2}},$$

или

$$l_{K'}^{K} = l_{K'}^{K'} \sqrt{1 - \frac{v^2}{c^2}}. \qquad (15)$$

Итак, мы знаем теперь, чему равна длина масштаба, неподвижного в системе K' и измеренного в системе K. Она равна длине,

измеренной в K', но умноженной на $\sqrt{1-\frac{v^2}{c^2}}$, т. е. она меньше длины, измеренной в K'.

Но мы можем поступить и наоборот. Пусть имеется масштаб, покоящийся в системе K. Его длина l_K^K. Мы хотим измерить его в системе K', по отношению к которой он движется. Если совершенно таким же способом мы произведем измерение, то получим

$$l_K^{K'} = l_K^K \sqrt{1-\frac{v^2}{c^2}}.$$

Итак, у нас есть два масштаба, которые первоначально (в системе K или K') были совершенно одинаковы. Мы передаем один масштаб в другую систему. Обозначим один из них, покоящийся в K, через A, а другой, покоящийся в K', через B. Тогда наблюдатель в K скажет: мой масштаб A покоится, а B движется. Значит, B согласно (15), короче, чем A. Наблюдатель в K' скажет: B покоится, а A движется, значит A короче, чем B. Получается такого рода ситуация, что относительно одного масштаба (A) говорят, что он короче, чем B, а с другой стороны, относительно этого же масштаба говорят, что он длиннее, чем B. Парадокс видели именно в том, что говорили: A одновременно и меньше B и больше B.

Есть ли здесь парадокс? Конечно, никакого парадокса нет, потому что эти высказывания — один раз, что A больше B, а другой раз меньше B — относятся к двум различным способам измерения: один раз измеряют при помощи средств (координат и часов), покоящихся в одной системе, другой раз — в другой, и поэтому совершенно не удивительно, что получают различные результаты.

Возьмем грубый пример. Двое встречаются на улице и проходят друг мимо друга. Один говорит: «Я прошел справа от вас, вы находились слева», а второй скажет: «Нет, это я прошел справа, а вы находились слева». Очевидно, это «противоречие» бессодержательно. Для меня *он* находился слева, а для него *я*. Точно так же и здесь. Нельзя усмотреть неприятностей в том, что, измеряя разными способами, мы получили различные результаты. Способ измерения связан с понятием одновременности, а так как одновременность различна, то и зарубки будут разные. Я не вижу здесь никакой трудности. Я должен сказать, что при той ясной формулировке, какая с самого начала была дана Эйнштейном, мне даже несколько трудно сформулировать, в чем видят парадокс. Если говорят о парадоксе, то это просто недоразумение. Можно не соглашаться с фактической стороной, но в логической постановке проблемы я не вижу никаких затруднений.

Заметим теперь следующее. Если я измерю мой масштаб, когда он находится в покое, а потом измерю его после того, как передам его в движущуюся систему, то он окажется короче; находясь в следующей, более быстро движущейся системе, он будет для меня еще короче; самая большая длина будет тогда, когда масштаб измеряется в системе, в которой он покоится. Эту длину называют *длиной покоя*.

И вот спрашивали так: но какова все-таки *действительная длина* масштаба? На это ответ может быть только один: это бесплодный вопрос. *Нет «действительной» длины* масштаба; в зависимости от условий измерения длина получится различной. Известен такой хороший пример. Если вы смотрите на масштаб, то угол, под которым вы его видите, с большего расстояния будет меньше. Имеет ли смысл спрашивать, какой *действительный угол* зрения соответствует этому масштабу? Это не имеет смысла: смотря по тому, на каком расстоянии вы находитесь, угол будет различный. Нет такого понятия, как «действительная» длина, т. е. длина безотносительно к системе отсчета. По определению, длина движущегося масштаба измеряется так-то, а длина неподвижного так-то, и нигде не сказано, что это всегда должно быть одно и то же. Спрашивать о «действительной» длине — значит не понимать того, как ставится задача.

Часто высказывается другого рода утверждение, на которое ответить нетрудно, но обычно в изложении теории относительности эта сторона недостаточно оттеняется. Говорят так: если я измеряю масштаб из K, то получаю одну длину, а если измеряю из K', то другую. Следовательно, это *субъективно*, т. е. длина масштаба — субъективное понятие и зависит от наблюдателя. Ничего не может быть неправильнее этого утверждения. Субъективности здесь нет никакой. Когда я говорю про то, что́ *видит наблюдатель* в K' и что́ *видит наблюдатель* в K, это просто удобный способ выражаться. Единственное, что важно,— это какой системой отсчета я пользуюсь. Если я говорю, что, пользуясь для масштаба, находящегося в покое в K, координатами и временем K', я получу укорочение, то сижу ли я сам в K или в K', это не играет роли. Можно было бы вообще устранить наблюдателей и пользоваться автоматическими приборами. Поэтому ни о какой субъективности не может быть и речи.

Наконец, самый неприятный вопрос, который также довольно часто ставится. У меня есть масштаб, имеющий определенную длину. Я его переношу в другую систему и получаю укорочение. Легко видеть, что это то же самое укорочение, которое постулировал Лоренц для объяснения опыта Майкельсона. Спрашивается, *действительное* ли это укорочение или *кажущееся*? На этот вопрос и в хорошей литературе получаешь, как это ни странно, различ-

ные ответы. Позвольте мне поэтому сказать несколько слов о том, насколько реальной вещью является лоренцово сокращение.

Рассуждают так: имеется движущийся стержень; мы измеряем его из неподвижной системы. Результат, который мы получаем (т. е. сокращение), зависит от того, как мы определили в нашей системе одновременность. Определили бы иначе, получили бы другой результат. Значит, это сокращение никакому реальному явлению не соответствует.

На это нужно возразить следующее: всегда и везде, при всех заведомо реальных явлениях может быть поставлен тот же вопрос. Пусть вы нагреваете стержень и он удлиняется. Тогда я говорю: если я измеряю длину нагретого стержня так, что приставляю к нему холодный масштаб, то получаю одну определенную длину, но если бы я дал другой способ измерения — приставьте масштаб и ждите, пока он тоже нагреется,— то я мог бы получить другое удлинение или даже укорочение, если масштаб сделан из материала с большим коэффициентом расширения.

Вы видите, что и здесь в зависимости от метода, которым вы измеряете, вы получаете различный результат. Для того чтобы сказать, чему равняется длина, нужно сказать, что такое длина, т. е. нужно указать метод измерения. В зависимости от метода будут и результаты различны. Но это не затрагивает вопроса о том, реально ли явление или нереально. Неправильность постановки вопроса заключается в том, что хотят высказать нечто о длине *одного* стержня, например удлиняется он или укорачивается, вне зависимости от других тел. Правильно сказать так: если у вас есть два покоящихся стержня, а потом один движется, то между длинами этих двух стержней получается определенное соотношение. То, что, измеряя указанным способом, одинаковым в K и K', вы получаете в обоих случаях укорочение,— это вполне реальный факт. Его трудно установить прямым опытом, потому что при самых больших скоростях, которые могут иметься в нашем распоряжении (скорость Земли), разница равна 1/200 000 000, но принципиально такой опыт возможен. Вопрос в том, даст ли он тот результат, который предсказывает теория относительности, или не даст. Вопрос этот касается взаимоотношения двух масштабов, движущихся друг относительно друга. Об этом взаимоотношении делается определенное высказывание, говорится, как проверить его на опыте, и утверждается, что опыт даст такой-то результат. Чего-либо более реального требовать нельзя: *опыт* может показать, так это или не так, и это наилучшее доказательство того, что вопрос реален. Если вы говорите, что у вас не 100 рублей, а 10 тысяч копеек, то это кажущееся увеличение вашего капитала, потому что оно не зависит ни от какого опыта, это вопрос чисто логический. Здесь же вопрос

не логический, здесь делается утверждение относительно взаимоотношения двух определенных стержней. Либо оно существует в природе, либо не существует. Можно утверждать, что его не существует; но если оно существует, то оно реально.

Можно сказать в качестве пояснения (хотя я считаю предыдущее рассуждение вполне убедительным) еще и так: пусть вы находитесь в системе K и имеете одинаковые стержни. Один из них вы передаете в другую систему, измеряете его и получаете другую длину. Вы говорите: можно было бы сделать так [1], чтобы длина переданного стержня осталась прежней. Вы передаете еще один стержень в третью систему и получаете опять другую величину. *Но как бы вы ни измеряли*, вы не сделаете так, чтобы длины всех (по-разному движущихся) стержней были одинаковы. Другими словами, если вы остаетесь все время при одном и том же методе измерения, то принцип относительности утверждает, что, измеряя масштаб, движущийся с одной скоростью, а потом измеряя масштаб, движущийся с другой скоростью, вы получите два различных значения длины. Можно говорить только о правильности или неправильности этого утверждения.

Наконец, говорили и так: можно считать, что взаимное укорочение наступает *без причины*; не видно, *почему* второй масштаб станет короче, если первый начнет быстрее двигаться. На это можно ответить следующее. Дело вовсе не в том, как изменяется второй масштаб, потому что в таком виде вопрос не имеет смысла. Нужно спросить, *по отношению к чему* изменяется? Что же касается взаимоотношения двух масштабов, то причина их взаимного сокращения в том, что они находятся друг по отношению к другу в движении. Ведь вопрос заключается именно во *взаимоотношении* масштабов, так что и с этой стороны все в порядке.

Теперь мы должны перейти ко второму основному измерению — измерению хода часов. Пусть имеются неподвижные в K' часы, и интервалом времени я буду называть разность между показаниями часов в один момент и в другой. Пусть я вижу стрелку сначала на t_1', а затем на t_2'. Тогда я говорю, что часы отсчитали интервал времени $\tau_{K'}^{K'} = t_2' - t_1'$ (индексы K' означают то же, что выше).

Теперь я хочу определить ход часов, неподвижных в K', из системы K. Как это сделать? Каков ход часов, движущихся по отношению ко мне? Ведь я не могу сравнить их с моими неподвижными часами, потому что движущиеся часы от меня удаляются. Наиболее разумное определение, которое мы всегда и даем, следующее: пусть движущиеся часы имеют в системе K' координату x'. Пусть в момент, когда они показывают t_1', они

[1] [Т. е. так определить длину движущегося стержня.]

как раз совпадают с часами в системе K, которые показывают t_1. Пусть *другие* часы в K, с которыми совпадают движущиеся, тогда, когда они показывают момент t'_2, показывают t_2 (это будут непременно другие часы, потому что за это время движущиеся часы переместились). Так вот, по определению, интервалом времени в K будет $t_2 - t_1$, т. е.

$$\tau_{K'}^{K} = (t_2 - t_1)_{x'_2 = x'_1 = x'}.$$

При помощи формул преобразования (14) легко вычислить, что

$$\tau_{K'}^{K} = \frac{\tau_{K'}^{K'}}{\sqrt{1 - \frac{v^2}{c^2}}}. \qquad (16)$$

Обратно, если имеются часы, неподвижные в K, и измеряется их ход из K', то получится

$$\tau_{K}^{K'} = \frac{\tau_{K}^{K}}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

О чем это говорит? Неподвижные часы дают большее время, т. е. ход часов, неподвижных в некоторой системе, если измерять

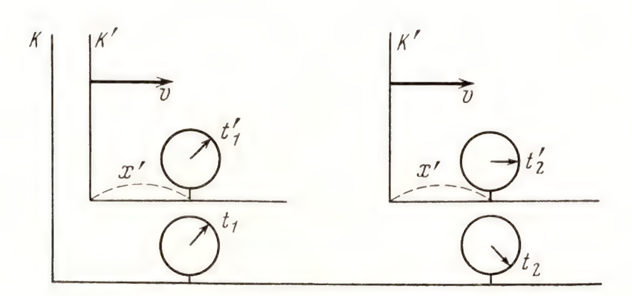


Рис. 47

из другой системы, оказывается более медленным. Но опять таки, если я буду проверять ход часов, покоящихся в этой второй системе, из первой, то я снова скажу, что движущиеся часы идут медленнее. Таким образом, с часами получается то же самое, что и с длиной. В зависимости от того, из какой системы они рассматриваются, ход их оказывается различным, причем часы идут медленнее, если они движутся относительно меня. Но это обратимо:

если A скажет, что часы B идут медленнее, то B скажет, что часы A идут медленнее. Здесь тот же «парадокс», но и здесь совершенно ясно, почему это происходит. Первый случай такой: я измеряю интервал времени по неподвижным часам. Второй случай: часы движутся, и я сравниваю их один раз с одними неподвижными часами, другой раз с другими. Это совершенно разные способы измерения, и то, что они дают различные результаты, совершенно естественно.

Таким образом, здесь можно повторить все, что было сказано раньше о длине стержня. Это вовсе не парадокс, а вопрос о том, каким «рецептом» я пользуюсь. Наконец, и вопрос о реальности решается совершенно так же: это замедление хода часов реально в той же степени, что и сокращение длин.

Однако мы можем здесь научиться еще кое-чему, что показывает, насколько это явление реально. Пусть в точке O двое часов находятся в покое. Затем одни часы начинают перемещаться, и, когда они пройдут расстояние x, мы их останавливаем (само собой разумеется, останавливаем не их ход, а их перемещение). Так как покоящиеся часы шли медленнее, то, когда двигавшиеся часы придут в конечную точку A, они будут показывать время t' иное, чем часы, оставшиеся в O, или, что то же самое, часы, покоившиеся в A. Если сделать простой расчет, то легко убедиться,

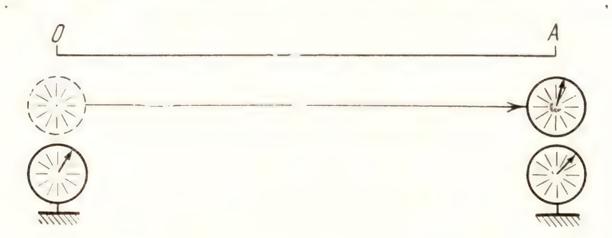


Рис. 48

что часы, передвинувшиеся в A, отстанут от часов, покоящихся в O (или A) на

$$t - t' = \frac{x}{v}\left(1 - \sqrt{1 - \frac{v^2}{c^2}}\right).$$

Не зная теории относительности, можно было бы определить синхронизм тем, что приписать удаленной точке время, показываемое часами, перенесенными в эту точку. Теперь мы видим, что если теория относительности права, если ее определение синхронизма правильно, то в зависимости от того, скорее или мед

леннее я переношу часы, они будут показывать различное время. Перенос часов дает не то же самое, что и световая синхронизация, и более того, дает различный результат в зависимости от скорости переноса, потому что разница в показаниях часов зависит от v и v^2. Значит, если теория относительности правильна, то пользоваться для синхронизации переносом часов нельзя. В одном только случае, а именно при $v \to 0$ и соответственно весьма длинном времени переноса, мы все-таки получим $t' \approx t$. Таким образом, мы можем получить путем переноса тот же результат, что и при определении одновременности по Эйнштейну в том случае, если будем осуществлять перенос бесконечно медленно, но только в этом случае. Строго говоря, если перевозят часы на пароходе и затем синхронизируют их при помощи радиосигналов, то это дает две различные установки часов. Хронометр, перевезенный на корабле, окажется не синхронным, если вы будете проверять его ход при помощи радиосигнала. Конечно, практически разница будет настолько ничтожна, что мы ее не обнаружим, но принципиально теория относительности утверждает, что при переносе не будет однозначного определения одновременности, потому что ход часов зависит от их скорости.

На этом, в сущности, можно было бы вопрос о времени закончить, если бы как раз с часами не возникла еще одна неприятность или еще один парадокс, который особенно много разбирался в литературе раньше, но к которому возвращаются еще и теперь.

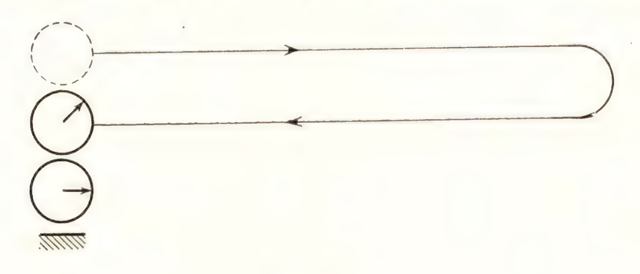


Рис. 49

Дело заключается в следующем. Представьте себе, что имеется двое часов, первоначально неподвижных в системе K. Затем одни часы начинают двигаться и удаляются на некоторое расстояние. При этом они несколько отстанут от покоящихся часов. Затем мы заставляем их двигаться обратно. Поскольку они опять движутся, они еще более отстают, и, следовательно, когда они вернутся в прежнюю точку, их показание будет меньше, чем у покоящихся часов. Итак, пусть часы U_1 находятся в какой-нибудь галилеевой системе, т. е. в любой системе, движущейся

по отношению к звездам равномерно и прямолинейно. Другие часы U_2 я несу с известной скоростью в определенном направлении, а затем несу их обратно на то же самое место. Часы U_2 должны отстать от часов U_1. Ну, что же, мы уже привыкли к таким историям. Но эта история хуже, и вот почему: в конце процесса я сравниваю часы, находящиеся в *одной и той же* точке. Все, что я теперь высказываю, не зависит от каких-либо определений одновременности, потому что относится к одной и той же точке. Но тогда возникает парадокс, который заключается в следующем. Я говорил, что я имею часы U_1, неподвижные по отношению к моей системе, и что я двигаю туда и обратно часы U_2. Но я мог бы сделать все совершенно так же в системе координат, по отношению к которой неподвижны часы U_2, а двигаются туда и обратно часы U_1. Я пришел бы в ту же точку, но теперь должны были бы идти замедленно часы U_1, т. е. U_1 должны были бы отстать от U_2. Значит, смотря по рассуждению, мы должны говорить, что отстали либо часы U_1, либо часы U_2. Это два различных реальных высказывания. Но ведь положение стрелок на обоих часах, находящихся в одном и том же месте, — это факт, для всех систем отсчета одинаковый. Значит, если я из одного рассуждения вывел, что U_1 будут показывать 12 часов, а U_2 — 12 ч. 10 м., из второго же рассуждения — что если U_1 показывают 12 часов, то U_2 показывают 11 ч. 50 м., то это действительно никуда не годится. Если бы дело обстояло так, то это было бы противоречием. Но здесь нет противоречия, потому что мы допустили ошибку.

В первом случае я говорил, что часы U_1 неподвижны в галилеевой системе, и все мое рассуждение было законно. Если же я беру систему, в которой неподвижны часы U_2, то эта система не галилеева. Ведь для того, чтобы вернуться обратно, часы U_2 должны были где-то изменить свою скорость, т. е. должны были иметь ускорение, а, значит, связанная с ними система отсчета уже не галилеева и требовать по отношению к ней того же, что имеет место в галилеевых системах, нельзя. Первоначально рассуждение было проведено для галилеевой системы: если U_1 покоятся в галилеевой системе и движутся U_2, то U_2 отстают. Если же U_2 будут в галилеевой системе, то это будет уже *другой* опыт, так как тогда U_1 не будут в галилеевой системе (иначе U_1 не смогут вернуться) и тогда U_1 отстанут. Здесь противоречия нет. Противоречие получилось только потому, что я считал эти две системы с точки зрения принципа относительности тождественными, а они не тождественны, так как одна из них наверняка не галилеева.

Что происходит в действительности, как идут ускоренно движущиеся часы и почему их ход изменяется, на это специальная

теория относительности ответить не может, ибо она вообще не занимается вопросом об ускоренно движущихся системах отсчета. До общей теории относительности было совершенно определенно выяснено, что противоречия здесь нет, но как подойти к решению вопроса, как, грубо говоря, объяснить явление, этого сказать нельзя было. Общая теория относительности без труда все объясняет. В ней естественным образом получается, что будут отставать те часы, которые находятся в негалилеевой системе. В круг ведения общей теории относительности входят как галилеевы, так и негалилеевы системы, и она может решить вопрос по существу. Но парадокса здесь, во всяком случае, нет. Если и сейчас в некоторых книгах ссылаются на то, что это рассуждение гибельно для специальной теории относительности, что оно ставит ее в тупик, то это просто неверно. Здесь от специальной теории относительности требуют того, чего она никогда не обещала и относительно чего она заранее говорит, что она этого не дает. Можно утверждать, что специальная теория относительности неверна, что природа построена иначе, но логического тупика здесь нет. Это, мне кажется, ясно, и излагать так, будто здесь есть тупик, значит неправильно указывать, в чем дело.

Очень интересными рассуждениями по этому поводу занимались многие. Особенно интересно этот вопрос разработал в очень красиво написанной статье Ланжевен [1]. На самое явление обратил внимание Эйнштейн.

Конечно, дело не сводится к тому, что запаздывают часы. *Всякий* процесс, в том числе и всякий периодический процесс, при перемещении будет запаздывать, т. е. все процессы будут протекать медленнее. Поскольку в атоме происходят периодические движения, эти периодические движения для атома, удалившегося и пришедшего обратно, запоздают по отношению к движениям в том атоме, который оставался неподвижным в галилеевой системе. И вот Ланжевен говорит: возьмите человека или какое-нибудь другое живое существо. Все процессы в нем также будут замедленны. Значит, если заставить его быстро двигаться, а затем вернуть обратно, то все процессы запоздают, этот человек будет медленно жить. Если сделать v почти равной скорости света, то разница может быть огромной: скажем, человек, который остается на месте, прожил 20 лет, а человек, который отправился и вернулся, прожил только 2 года. Первый может состариться, а тот, который двигался туда и обратно, будет выглядеть молодым человеком.

Я не думаю, чтобы физики отказались от этого вывода. Физики вообще (многие физики) неохотно вступают на путь биологии или

---

[1] [*P. Langevin.* Scientia, 10, 31, 1911.]

психологии. Но нет, в сущности, никакого основания думать, что здесь будет иначе. Несомненно, что организм подчинен основным физическим законам. Человек падает на землю так же, как и свинец, живой он или не живой. Временные процессы протекают так же, как и ход часов. Этим вовсе не исчерпывается все; наоборот, я думаю, что банально говорить, будто физические законы все объясняют в жизни организма, но утверждение о быстроте протекания процессов остается справедливым и для живого организма.

Конечно, для заметных эффектов нужны колоссальные скорости. Тем не менее здесь мы снова стоим перед принципиально проверяемым фактом. Теория относительности утверждает, что это будет так. Можно сказать: «Попробуйте, проверьте это», но я думаю, что большинство релятивистов согласилось бы с таким выводом из теории, а не заявило бы: «Я знать не знаю и ведать не ведаю, что происходит в организме».

Для развлечения Ланжевен взял очень большие скорости. Для того чтобы получилось замедление в 100 раз, нужно, чтобы v отличалось от скорости света на 1/20 000. Тогда происходят невероятные вещи: здесь человек проживет 100 лет, а человек, который путешествовал туда и обратно, проживет всего лишь год. Ланжевен вычисляет, однако, какая нужна энергия для того, чтобы развить такую скорость. Оказывается, что раскачивать путешественника придется примерно в течение года и нужно для этого 400 миллиардов лошадиных сил.

Итак, мы приходим к заключению, что в рассмотренных вопросах эйнштейновской кинематики никаких логических противоречий нет. Есть вполне определенные утверждения относительно поведения реальных предметов. Нам остается рассмотреть еще один вопрос, на котором мы и закончим эйнштейновскую кинематику,— вопрос о сложении скоростей.

ОДИННАДЦАТАЯ ЛЕКЦИЯ

(4.IV 1934 г.)

Краткое резюме. Существуют ли абсолютные величины? Интервал. Что означает сложение скоростей. Теорема Эйнштейна о сложении скоростей. Следствия из этой теоремы. Скорости, превышающие с. Коэффициент увлечения. Сопоставление классической и релятивистской точек зрения.

В прошлый раз мы подробно остановились на некоторых следствиях, вытекающих из преобразования, связывающего координаты одной галилеевой системы с координатами другой, так называемого лоренцова преобразования. Мы поставили ряд вопросов, выяснение которых необходимо для дальнейшего, для того чтобы разобраться в других вопросах, которые, в сущности, нас больше всего и интересуют,— вопросах механики, электродинамики и т. д.

Нам встретился такой вопрос: найти длину движущегося масштаба. В одной из систем масштаб неподвижен. Как измерить его в движущейся системе? Для этого нужно дать определение длины движущегося масштаба. Мы дали такое определение и нашли, что длина масштаба, покоящегося в K' и измеренного в K' $(l_{K'}^{K'})$, связана с его длиной, измеренной в системе K $(l_{K'}^{K})$, соотношением

$$l_{K'}^{K} = l_{K'}^{K'}\sqrt{1-\beta^2}\,, \qquad \left(\beta = \frac{v}{c}\right).$$

Обратно, если мы хотим измерить из K' масштаб, покоящийся в системе K, то мы опять получим укорочение, т. е.

$$l_{K}^{K'} = l_{K}^{K}\sqrt{1-\beta^2}.$$

На первый взгляд получается, что одна и та же величина может быть и длиннее и короче другой, что, конечно, абсурдно. Но дело в том, что здесь два различных процесса измерения, и поэтому не удивительно, что результаты различны.

То же самое мы получили и при сравнении хода часов: движущиеся часы идут медленнее, чем покоящиеся, какую бы систему координат мы ни рассматривали. В случае часов особенно ясно, что никакой парадоксальности нет. Пусть имеются системы K и K' и мы хотим измерить ход часов, покоящихся в K', из системы K. Что это значит? Пусть часы A и B показывают 12 (рис. 50). Часы B движутся, через некоторое время приходят в точку B_1 и показывают в этот момент 1 час. Что будут показывать в этот же момент часы в точке A_1? Так как часы B идут медленнее, то на A_1 будет, скажем, 1 ч. 5 м.

Обратно, что означает измерение хода часов, покоящихся в K, с точки зрения K'? Я должен взять часы A и сравнить их показания один раз с часами B, другой раз с часами B_2, против которых окажутся часы A, передвинувшись в положение A_2. Теперь на часах A (в точке A_2) будет 1 час, а на B_2 — 1 ч. 5 м. Вы видите, что здесь два различных процесса измерения, а потому и не удивительно, что получаются такие результаты. Таким образом, отсчет времени и отсчеты координат относительны, т. е. если происходит какое-нибудь событие, то мы не можем сказать, что оно имеет какие-то определенные координаты «вообще». Координаты x, y, z, t зависят от того, в какой системе мы наблюдаем это событие.

Относительно координат прежняя физика признавала то же самое. Конечно, если рассматривать явление в двух различных системах, то в одной из них координаты будут иметь одни значения, в другой — другие. Но считалось незыблемым, что показания часов во всех системах будут всегда одни и те же. Теория относительности разрушает это представление и говорит о времени то же, что и о пространстве.

Здесь есть, однако, один момент, который, может быть, недостаточно подчеркивается: что же при этом неотносительно, что же остается абсолютным? В общий разбор вопроса о том, что называть абсолютным и что относительным, мы вдаваться не будем,

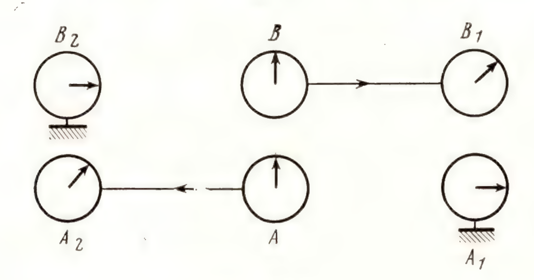


Рис. 50

а рассмотрим лишь то, что для нас нужно. Мы скажем так: если какая-нибудь величина, относящаяся к данному событию, будет иметь в разных галилеевых системах различное значение, то мы назовем эту величину относительной, если же она имеет во всех галилеевых системах одно и то же значение, если она инвариантна, то мы будем говорить, что для рассматриваемого класса систем это абсолютная величина.

Возьмем чистую геометрию — без всякого движения. Если имеется какая-то точка, то ее координаты x, y, z относительны, их

значения различны, если смотреть на эту точку из разных координатных систем. Мы скажем, что понятие координат в геометрии относительно и всецело зависит от системы отсчета, которой мы пользуемся. Если имеются две точки, то координаты каждой из них относительны, но расстояние между двумя точками $\sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}$ (где $\Delta x = x_2 - x_1$, $\Delta y = y_2 - y_1$, $\Delta z = z_2 - z_1$) не зависит от того, из какой (прямоугольной) системы его рассматривать. Эта величина для всех прямоугольных систем инвариантна, и мы скажем, что для этих систем она абсолютна. Наше рассуждение относится пока к числам, к буквам. Но мы связываем с этими числами определенную реальность. Мы говорим: координаты можно измерить и получить числа как результаты действительных измерений. Это будут относительные величины. Но мы знаем, что, например, длина стержня остается инвариантной, не положение стержня, а именно его длина, т. е. расстояние между его концами.

Спрашивается, при переходе к классу инерциальных систем, т. е. при рассмотрении не только поворотов координатных осей, но и движущихся галилеевых систем, что оставалось абсолютным в прежней физике? Абсолютным оставался корень из суммы трех квадратов, который мы называем расстоянием. Кроме того, абсолютным оставалось значение времени или разность времен для двух событий. Мы имели, таким образом, две абсолютные величины.

Теория относительности это отрицает: $t_2 - t_1$ не есть абсолютная величина; при переходе от системы к системе она меняется, абсолютным же остается выражение $(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - c^2 (\Delta t)^2$. Эта величина для всех галилеевых систем неизменна, инвариантна, т. е., по нашему определению, она абсолютна. Корень квадратный из этого выражения обычно принято называть *интервалом*.

Возникает вопрос: какой физический процесс или предмет связан с этой величиной? Мы видели, что расстояние представлялось при помощи стержня. Что же является представителем интервала, абсолютно инвариантного для всех галилеевых систем? Об этом мы поговорим позже. Пока же я замечу только следующее: совершенно напрасно думают, будто принцип относительности разрушил понятие абсолютного (в указанном смысле). Он только передвинул это понятие. Правда, в прежней физике интервал также был абсолютен, но важно то, что абсолютны были и расстояние само по себе и t само по себе. Принцип относительности это изменил, но не отказался от абсолютного. Это очень существенный момент, потому что в физике нас, конечно, прежде всего интересуют соотношения именно между такими абсолютными величинами. Вся евклидова геометрия, в сущности, строится на том, что рас-

стояние есть абсолютная величина. Все учение о треугольнике исходит из этого же основного факта. Мы еще будем об этом говорить, и то обстоятельство, что в теории относительности также есть абсолютная величина (интервал), мы еще используем.

Таковы результаты, относящиеся к кинематике Эйнштейна, к которым мы пришли и которые существенным образом отличаются от прежней кинематики. Но, я думаю, вы сами теперь уверены в том, что ничего сверхъестественного здесь нет. В основу релятивистской кинематики кладутся утверждения относительно поведения твердых тел (и вообще тел) по отношению к свету.

Напрасно думали, что классика этого не делала. Она также полагала в основу определенные утверждения, но только другие. Опыты, которые мы имеем, говорят о том, что все подтверждает поведение тел, постулируемое теорией относительности. Конечно, есть еще ряд экспериментально нерешенных вопросов; разница в ряде случаев столь мала, что невозможно произвести опыты, которые дали бы дальнейшее подтверждение, но то, что мы уже знаем, свидетельствует, что утверждения теории относительности не только не ведут к противоречиям, но, наоборот, находятся в согласии с известными опытами в противовес утверждениям классики, которые не давали такого согласия. Например, опыт Майкельсона противоречит утверждениям прежней кинематики. Насколько мы знаем, опыты пока говорят за то, что постулаты теории относительности, касающиеся реальных физических тел, оправдываются, между тем как о прежних положениях этого сказать нельзя.

Теперь мы перейдем к последнему вопросу, которым кинематика Эйнштейна будет исчерпана. Я имею в виду вопрос о скоростях. Вы знаете, как складываются скорости в прежней, классической теории. Если тело имеет сразу две скорости, то говорят, что это то же самое, как будто тело обладает одной скоростью, равной $w = u + v$, причем эту сумму нужно понимать векторно.

Это чрезвычайно неясная формулировка. Что вообще значит выражение «тело имеет две скорости»? По существу это выражение не имеет смысла: если тело как-то движется, то оно имеет одну скорость, определенную и по направлению и по величине. Здесь, как во всех такого рода вещах, нужно сначала уточнить, что мы будем понимать под утверждением, что тело имеет две скорости.

Если посмотреть, как пользуются сложением скоростей в физике, то, по существу, здесь имеют в виду два различных понятия сложения, каждое из которых подчиняется своему определению. Первое — это построение, вытекающее из определения понятия скорости, которое говорит: если тело движется и имеет некоторую скорость u, то за известный промежуток времени оно проходит расстояние $u\Delta t$ (мы будем говорить только о скоростях, постоян-

ных по направлению и величине, так как к общему случаю можно перейти путем предельного перехода). Представим себе, что после этого тело вдруг изменило направление движения и в течение такого же промежутка времени Δt прошло расстояние $v\Delta t$. Мы можем спросить: какую скорость (по величине и направлению) должно иметь тело, чтобы за такой же промежуток времени Δt сразу (прямолинейно) прийти из начальной точки в конечную? Тогда из построения (рис. 51) ясно, что $\mathbf{w} = \mathbf{u} + \mathbf{v}$. Мы говорим, что эта скорость $\mathbf{w}$, при которой тело за такой же промежуток времени приходит в ту же конечную точку, по определению, есть сумма скоростей $\mathbf{u}$ и $\mathbf{v}$, что наличие скорости $\mathbf{w}$ и означает, что тело имеет одновременно две скорости: $\mathbf{u}$ и $\mathbf{v}$. Это определение. Им пользуются, когда разлагают скорость на слагающие. Это оправдано и целесообразно, потому что мы считали и считаем, что законы механики придают этому физический смысл.

Возьмем самый простой пример. Представьте себе, что в точке A находится какое-то тело (рис. 52). Вы даете ему толчок, оно приобретает известную скорость и через некоторое время приходит в точку B, где вы его останавливаете. Вы снова даете толчок, и оно, двигаясь с определенной скоростью, приходит в точку C. Тогда, основываясь на опыте, вы утверждаете, что если вы сразу дадите оба толчка, выстрелите сразу с двух сторон, то тело пойдет именно по AC, скорость его будет слагаться из обеих скоростей. Таким образом, в том, что параллелограмм скоростей справедлив,

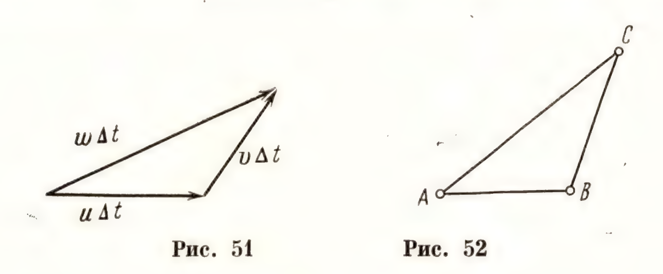


Рис. 51 Рис. 52

принцип относительности ни при чем. Все сказанное относится к какой-то одной определенной системе отсчета, в которой измеряются и $\mathbf{u}$, и $\mathbf{v}$, и $\mathbf{w}$, и сумма двух скоростей в *этом* смысле находится по правилу параллелограмма в новой кинематике так же, как и в прежней.

Но часто, когда говорят о сложении скоростей, имеют в виду другое, и это понятно, потому что целый ряд задач требует другого подхода. Вот что я имею в виду и что нас специально интересует там, где речь идет о движущихся средах. Самый простой случай,

который излагается во всех учебниках, следующий. Поезд имеет по отношению к земле определенную скорость. В вагоне поезда двигается тело, так что по отношению к поезду оно имеет определенную по величине и направлению скорость. Часами и масштабами, покоящимися относительно поезда, вы измеряете скорость тела по отношению к поезду. Пусть величина этой скорости u. Но движение тела в поезде может наблюдать и человек, находящийся на земле. По отношению к земле это тело также имеет определенную скорость. Пусть человек, находящийся на земле, измерил *своими* масштабами и *своими* часами скорость поезда по отношению к нему. Пусть эта скорость равна v. Спрашивается, если известны u и v (в механике v называется обычно переносной скоростью, а u — относительной скоростью), то чему равняется абсолютная скорость тела w? В прежней классической физике ответ был таков: она равна сумме этих двух скоростей, т. е. $w = u + v$.

Вы видите, что здесь совершенно другая постановка вопроса. Первая скорость v — это скорость, измеренная в одной системе, K; скорость u — это скорость, измеренная в другой системе, K'. Спрашивается, какова будет скорость w, измеренная опять-таки в первой системе? Ясно, что здесь ответ на вопрос о сумме скоростей всецело будет зависеть от связи, которая имеется между метриками обеих систем — K и K', в частности от понятия одновременности. Поэтому нигде не сказано, что здесь тоже будет правило параллелограмма. Нам нужно теперь рассмотреть эту задачу снова, с точки зрения кинематики Эйнштейна, и вовсе не известно заранее, получим ли мы то же самое.

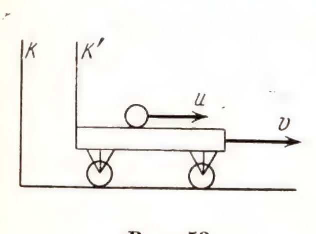


Рис. 53

Постановка вопроса совершенно ясна: в нашей системе K' движется тело, т. е. x', y', z' являются функциями от t'. Скорость его мы назовем u', причем $u'_x = \frac{dx'}{dt'}$, и аналогично для u'_y и u'_z. Что мы назовем скоростью по отношению к системе K? Здесь, по определению, мы имеем $u_x = \frac{dx}{dt}$, $u_y = \frac{dy}{dt}$, $u_z = \frac{dz}{dt}$. Мы знаем, что x, y, z, t находятся в определенной зависимости от x', y', z', t'. Значит, если x', y', z' — функции от t', а x, y, z — функции от x', y', z', t', то и $\frac{dx}{dt}$, $\frac{dy}{dt}$, $\frac{dz}{dt}$ будут функциями от t'. Нам нужно теперь определить $\frac{dx}{dt}$, $\frac{dy}{dt}$, $\frac{dz}{dt}$, которые и представляют собой скорость в системе K.

Формулы преобразования у нас таковы:

$$x' = \frac{x - vt}{\sqrt{1-\beta^2}}, \qquad y' = y, \qquad z' = z, \qquad t' = \frac{t - \frac{vx}{c^2}}{\sqrt{1-\beta^2}}.$$

Для того чтобы найти, как связаны одни производные с другими, нужно продифференцировать

$$\frac{dx'}{dt'} = \frac{\frac{dx}{dt} - v}{\sqrt{1-\beta^2}} \frac{dt}{dt'}, \qquad \frac{dy'}{dt'} = \frac{dy}{dt} \frac{dt}{dt'}, \qquad \frac{dz'}{dt'} = \frac{dz}{dt} \frac{dt}{dt'}.$$

Остается найти еще $\frac{dt}{dt'}$. Выразив t через t', находим

$$\frac{dt}{dt'} = \frac{1 + \frac{v}{c^2} \frac{dx'}{dt'}}{\sqrt{1-\beta^2}}.$$

Из этих четырех уравнений мы без труда можем получить решение интересующего нас вопроса. Здесь простая арифметика и, выполнив этот несложный расчет, мы получим

$$u_x = \frac{u'_x + v}{1 + \frac{vu'_x}{c^2}}, \qquad u_y = \frac{u'_y \sqrt{1-\beta^2}}{1 + \frac{vu'_x}{c^2}}, \qquad u_z = \frac{u'_z \sqrt{1-\beta^2}}{1 + \frac{vu'_x}{c^2}}. \qquad (17)$$

Итак, если в вагоне тело имеет скорость u'_x, u'_y, u'_z и если вагон движется по отношению к земле со скоростью v, то полученные формулы показывают, с какой скоростью тело движется по отношению к земле.

Можно задать обратный вопрос — разрешить эти уравнения относительно штрихованных величин. Если принцип относительности справедлив, то ответ должен быть равносилен таким же формулам, только с заменой v на $-v$. Нетрудно убедиться прямым расчетом, что это так и есть.

Полученные формулы выражают знаменитую теорему Эйнштейна о сложении скоростей. Вы видите, что они отличны от того, что дает классика. Действительно, что давала классика? Классика говорила, что u есть векторная сумма переносной и относительной скоростей. Переносная скорость имеет в наших осях всего одну компоненту. Значит, по классической кинематике $u_x = u'_x + v$, $u_y = u'_y$, $u_z = u'_z$. Здесь же вы видите совершенно другие соотношения.

Мы разложили скорость $\mathbf{u}$ на компоненты. Можно спросить, чему равна абсолютная величина скорости. *Этот* вопрос решается по-прежнему — по правилу параллелограмма: возьмите компоненты, возведите в квадрат и сложите. Для дальнейшей дискуссии удобно при этом ввести угол между $\mathbf{u}'$ и осью x в системе K'. Назовем его α'. Так как мы принимаем старую евклидову геометрию, то $\operatorname{tg} \alpha' = \frac{\sqrt{u_y'^2 + u_z'^2}}{u_x'}$. Отсюда можно определить $\cos \alpha'$, и в результате для u^2 получим

$$u^2 = \frac{u'^2 + v^2 + 2u'v \cos \alpha' - \left(\frac{u'v}{c^2} \sin \alpha'\right)^2}{\left(1 + \frac{u'v}{c^2} \cos \alpha'\right)^2}. \tag{18}$$

Это эйнштейновская теорема для абсолютной величины скорости, являющаяся, конечно, просто следствием выражений (17). Я подчеркиваю, что в теореме Эйнштейна о сложении скоростей уже не содержится никаких новых предпосылок. Она непосредственно следует из лоренцова преобразования, поскольку мы ясно сказали, какую задачу мы решаем, т. е. что мы понимаем под сложением скоростей. Перейдем теперь к следствиям теоремы Эйнштейна.

Предположим прежде всего, что как v, так и u очень малы. Что значит очень малы? Это значит, что квадраты этих величин и их произведения малы по сравнению с c^2 и ими можно пренебречь. Тогда эйнштейновские формулы обращаются в обычные формулы параллелограмма скоростей. Значит, если речь идет о малых скоростях, то с достаточным приближением все происходит по-прежнему. Движетесь ли вы в поезде, или даже стреляете из пушки в поезде, вы можете складывать скорости совершенно так же, как и раньше. Если вы учтете, что скорость света равна 300 тыс. *км* в секунду, то, конечно, никакой подобный опыт не сможет вам показать, имеется отличие от классики или нет. Но есть такие случаи, например при движении электронов, когда скорость близко подходит к скорости света, и там между старой и эйнштейновской кинематикой получается резкая разница. Для того чтобы все было более наглядно, представим себе, что оси расположены по-прежнему и мы рассматриваем движение только вдоль оси x. Это не вносит никаких принципиальных изменений, только формулы более просты. Мы получаем тогда

$$u = \frac{u' + v}{1 + \frac{u'v}{c^2}}. \tag{19}$$

Пусть теперь u' равна $^3/_4$ скорости света и v также составляет $^3/_4$ скорости света. Что мы сказали бы раньше? Мы сказали бы, что если вагон идет со скоростью $^3/_4$ световой и в вагоне движется тело тоже со скоростью $^3/_4$ световой, то по отношению к неподвижной земле тело имеет скорость, равную 1,5 скорости света. Что говорит Эйнштейн? Формула его такова, что если к скорости, меньшей скорости света, вы прибавляете скорость, тоже меньшую скорости света, то результирующая всегда будет тоже меньше скорости света. Можно сделать «слагаемые» равными 0,99 скорости света и все-таки «сумма» будет меньше c.

Итак, мы приходим к следующему следствию из этой теоремы: сложение двух скоростей, из которых каждая меньше скорости света, всегда дает скорость, меньшую скорости света.

А что получится, если одна из скоростей будет равна скорости света, т. е. если, скажем, в вагоне, имеющем скорость, меньшую скорости света, распространяется луч света, который имеет скорость c? Для того чтобы узнать какую скорость этот свет имеет по отношению к неподвижной земле, нужно сложить v и c. И вот, полагая $u' = c$, мы получаем из (19), что результат всегда равняется c, т. е. если к скорости, равной скорости света, прибавить скорость, меньшую c, то всегда получится скорость, равная c. Если бы это было не так, то это противоречило бы всему, что мы говорили раньше, так как мы исходили из того, что, какую бы систему отсчета мы ни взяли, скорость света в ней всегда одна и та же. Это был наш постулат. Но мы всегда можем рассматривать скорость света в неподвижной системе (земля) как сумму скоростей света в движущейся (в вагоне) плюс скорость этой последней (самого вагона). И правильно приложенная теорема сложения скоростей дает правильный ответ.

Эта теорема может быть применена и тогда, когда $u' > c$. Скорость v не может быть больше скорости света, так как v — это скорость движения нашей материальной системы K' по отношению к K. Мы знаем, что при $v > c$ формулы преобразования теряют смысл. Скорость тела, как такового, не может превысить скорость света. Но процесс может идти с любой скоростью. Если у меня в какой-то точке x_1 что-то случается в момент t_1, а в точке x_2 — в момент t_2, то я просто называю скоростью величину $\frac{x_2 - x_1}{t_2 - t_1}$. Мы видели, что существуют скорости, например фазовая скорость, которые могут превышать скорость света, и мы выяснили, что это не противоречит принципу относительности. Нельзя было только допустить, чтобы это была скорость сигнала, при помощи которого можно воздействовать, который мог бы быть причиной, вызывающей следствие; но понятие скорости может быть определено и для таких явлений, у которых скорость оказывается превышающей

скорость света. Все рассуждение, давшее нам формулу (19), ничуть не нарушится, если u' будет относиться именно к такому явлению. Что тогда получится?

Если в одной системе $u' > c$, то легко видеть, что и в другой системе всегда будет $u > c$. Таким образом, если к скорости, меньшей c, прибавить скорость, бо́льшую c, то в любой системе получится скорость, бо́льшая c. Я подчеркиваю, что это отнюдь не пустые разговоры, если говорят о скорости u в таких случаях, когда она оказывается больше c. Я хотел бы обратить ваше внимание на подобные случаи там, где с ними всегда приходится иметь дело,— в квантовой механике.

Позвольте мне поэтому остановиться на этом вопросе. Дело заключается в следующем (я не буду останавливаться на вопросах квантовой механики; многие из вас знают то, что мне нужно). Возьмем частицу — свободно движущийся электрон. Вы знаете, что поведение электрона описывается так называемой волновой функцией. Волновая функция для этого простого случая выражается так:

$$\psi = e^{i(kx'+\omega t')},$$

т. е. волновая функция есть не что иное, как плоская волна. Как всякая волна, она имеет фазовую скорость, и эта скорость есть $u' = \frac{\omega}{k}$, где $k = \frac{2\pi}{\lambda}$. Наряду с этим имеется определенная скорость у самой частицы. Назовем ее w'. И вот волновая механика дает определенную зависимость между u' и w'. Она указывает, что $\omega = \frac{2\pi\, mc^2}{h}$ (частота есть полная энергия, деленная на постоянную Планка), а $k = \frac{2\pi}{\lambda}$, где $\lambda = \frac{h}{mw'}$ (w' — скорость движения частицы). Если разделить ω на k, то получается

$$u' = \frac{c^2}{w'}. \tag{20}$$

Значит, уже в таком простом процессе, как движение частицы без силового поля, с точки зрения квантовой механики есть две скорости, связанные друг с другом. Одна скорость — это фазовая скорость ψ-волны, а другая — скорость движения материальной частицы. Они связаны зависимостью (20), так что фазовая скорость *всегда больше скорости света.*

Если мы перейдем теперь от системы K' к системе K, то обе скорости изменятся: u' будет другой и w' тоже будет другой. А как будет с (20)? Спрашивается: сохраняется ли это соотношение между двумя скоростями при переходе к другой системе отсчета? Переходя к другой системе при помощи эйнштейновских преобразований,

легко убедиться, что если $u'w' = c^2$, то и $uw = c^2$, т. е. эта зависимость, очень типичная для квантовой механики, инвариантна при лоренцовом преобразовании. При преобразовании Галилея она не была бы инвариантной. Кроме того, теперь обратили внимание и на то, что это единственное соотношение и другого инвариантного соотношения быть не может. Ясно, почему это так. Если бы было другое инвариантное соотношение, независимое от (20), то это значило бы, что две величины, u и w, были бы однозначно определены и они не могли бы изменяться, а между тем мы знаем, что они меняются.

Таким образом, можно сказать определенно, что рассуждения о скоростях, бóльших скорости света, это не просто интересный казус, а с этим действительно приходится иметь дело.

Позвольте дать еще один простой пример, который достаточно ясно показывает, я бы сказал, мощь этих взглядов, мощь эйнштейновской кинематики. Речь идет о френелевском коэффициенте увлечения, о котором мы много говорили раньше. Вы знаете, что коэффициент увлечения Френеля выражает тот факт, что, скажем, вода, которая имеет скорость, увлекает за собой свет не сполна. Если я все измеряю в одной системе K и получаю некоторую скорость света u_1 в покоящейся воде, а затем начинаю воду двигать со скоростью v, то измеренная в K скорость света в воде будет не $u_1 + v$, а $u_1 + v\left(1 - \frac{1}{n^2}\right)$, где n — показатель преломления.

Таким образом, скорости складываются не просто, а так, будто только часть воды движется со скоростью v. Это и есть то, что мы называем частичным увлечением по Френелю. Вы помните еще из исторического введения, какую важную роль играет при построении теории этот факт, и помните, что лоренцовская теория давала ему объяснение. У Лоренца выходило, что скорость света по отношению к воде равна $u_1 - \frac{v}{n^2}$, т. е. если вода движется, то в ней самой благодаря тому, что она движется по отношению к эфиру со скоростью v, скорость света изменилась. Когда же вы с помощью галилеева преобразования переходите к системе, неподвижной относительно эфира, то и получается известное нам выражение френелевского коэффициента.

Если посмотреть на всю проблему с точки зрения Эйнштейна, то ее можно решить очень просто (я нарочно беру среду, в которой нет дисперсии, т. е. нет зависимости скорости света от частоты; такой реальной среды, строго говоря, не существует, но с большим приближением это можно принять, это не меняет дела в принципе). По Эйнштейну, выходит так: возьмем какой-то световой импульс, распространяющийся в воде со скоростью c/n. Пусть

это имеет место тогда, когда вода находится в покое в системе K. Пусть теперь вода движется со скоростью v и, значит, покоится в системе K'. Принцип относительности говорит, что в этой системе скорость импульса должна быть такой же, какой она была в системе K, потому что здесь ставится тот же опыт. Следовательно, $u' = \frac{c}{n}$. Это уже отлично от результата, который давал Лоренц, так как у Лоренца было $u' - \frac{v}{n^2}$. Правда, обнаружить это экспериментально до сих пор нельзя, но по существу здесь Лоренц и Эйнштейн расходятся.

Спрашивается, что же мы получим с точки зрения Эйнштейна в системе K? Ответить на это очень просто. Используем теорему сложения. Если $u' = \frac{c}{n}$, то, чтобы узнать, чему равняется u, я пишу: $u = \frac{u' + v}{1 + \frac{u'v}{c^2}}$ — скорости не просто складываются алгебраически или геометрически, а привходит еще один фактор, наличие которого вытекает из метрики. Так как мы все делаем для первого порядка величины $\frac{v}{c}$, то можно написать $u = (u' + v)\left(1 - \frac{u'v}{c^2}\right)$, и мы получаем

$$u = u' + v\left(1 - \frac{u'^2}{c^2}\right) = u' + v\left(1 - \frac{1}{n^2}\right),$$

где просто, по определению, n есть коэффициент преломления. Итак, мы получаем эту формулу без дальнейшего, просто как результат новой кинематики. Причина, из-за которой получается «частичное увлечение», здесь, по существу, другая — не в том механизме, в котором ее видел Лоренц. Конечно, результаты в первом порядке должны быть одни и те же, так как свет идет туда и обратно; это более или менее ясно, и мы увидим далее, что лоренцовские уравнения для движущихся тел получаются из эйнштейновской теории. Но у Лоренца все получается потому, что у него в системе K' должна быть другая скорость света в воде и правильный результат вытекает из его теории известным образом случайно. Между тем мы видим теперь, что это результат чисто метрических свойств, т. е. чистая кинематика.

На этом кинематика заканчивается, и в следующий раз мы перейдем к более сложным вопросам применения всего учения Эйнштейна о времени и пространстве к установлению законов механики и электродинамики.

Позвольте в заключение еще раз в двух словах сопоставить классику и эйнштейновскую точку зрения.

Что дает классика? Классика утверждает: есть одна определенная система — система неподвижных звезд. Эта система K выделена, и никакого принципа относительности мы не знаем. В этой системе мы устанавливаем масштабы и часы, причем часы синхронизированы с помощью света. Всякий классик согласится с тем, что в системе K такой способ должен дать то же самое, что у Эйнштейна, потому что признается, что в этой системе свет распространяется во все стороны одинаково. Правда, классика говорила, что то же самое можно получить переносом часов,— это очень важный момент, но она согласилась бы, что в отношении света дело обстоит одинаково.

Классика далее говорит: если у вас есть система K', движущаяся по отношению к исходной системе K, то в ней мы устанавливаем метрику следующим образом: мы требуем, по определению, чтобы часы, находящиеся в K', показывали то же самое время, которое показывают часы K, с которыми часы K' встречаются при движении. Это наше определение одновременности в K' — мы требуем, чтобы $t' = t$. Затем мы считаем, что $x' = \alpha(v)(x - vt)$. Мы будем требовать, кроме того, чтобы один и тот же масштаб, измеренный из K и из K', давал одну и ту же длину. Тогда $x' = x - vt$. Устанавливать все таким способом было вполне правомерно. Но что́ совершенно недопустимо при построении физики? Классик утверждал дальше следующее: если вы передадите часы из K в K', то двое часов — оставшиеся в K и перенесенные в K',— которые раньше шли в системе K совершенно одинаково, будут показывать именно то время, которое вы установили. Классик утверждал, что именно так оно и будет. То же самое и с масштабами. Он утверждал что если мы передадим реальный масштаб, то получим для двух масштабов — оставшегося и переданного — одинаковую длину. Итак, классик давал определение, это было его право; но этого было недостаточно, и он, сверх того, утверждал определенные вещи относительно поведения масштабов и часов.

Что сделал Лоренц для объяснения опыта Майкельсона? Относительно времени он давал те же определения, что и классик, и подразумевал то же утверждение, что часы будут этому подчиняться. Но насчет масштабов он говорил: я утверждаю, что с реальными масштабами будет происходить укорочение, так что $x' = \dfrac{x - vt}{\sqrt{1-\beta^2}}$. С тем, что масштабы будут себя так вести, согласна и теория относительности, но Лоренц, кроме того, постулировал, что часы будут идти одинаково.

Наконец, теория относительности утверждает, что нужно установить понятия координат и времени так, как это делает Эйнштейн,— это также его право. Но фактическое утверждение теории относительности заключается в том, что реальные часы и мас-

штабы будут вести себя в соответствии с принципом относительности и принципом независимости скорости света от движения источника.

Значит, у всякой теории должны быть определения и нельзя сказать заранее, какое из них лучше. В этом отношении теория относительности не пошла ни на иоту дальше. Но этим еще ничего не сделано. Кроме этого, нужны утверждения относительно того, как будут себя вести реальные часы и реальные масштабы.

Таким образом, дело обстоит не так, что прежде не нужно было ничего определять, а теории относительности это нужно. Нет. Раньше утверждали одно, а теория относительности утверждает другое. В общепознавательном смысле все элементы, которые есть здесь, были также и там. Весь вопрос о том, кто прав *фактически*. Теория относительности утверждает, что она права, и, как сказано, опыты, по-видимому, дают ей право так говорить. Вопрос о том, справедлива ли теория относительности или нет,— это вопрос не логический, а фактический, это вопрос опыта, и только с этой точки зрения можно с теорией относительности соглашаться или не соглашаться.

ДВЕНАДЦАТАЯ ЛЕКЦИЯ

(22. IV 1934 г.)

Дополнительные замечания о лоренцовом преобразовании. Одновременные и одноместные пары событий в релятивистской и в классической кинематике. Более общий вид лоренцова преобразования. Вопросы о четырехмерности. Арифметизация геометрии и геометризация аналитики. Что именно нового внесла теория относительности. Прежние и новые инварианты. Времени-подобный и пространственно-подобный интервал. Вещественные представители интервала

Прошлый раз мы закончили вопрос о сложении скоростей в теории относительности. Позвольте сегодня сделать некоторые дополнительные замечания о лоренцовом преобразовании, которые несколько обобщат и уточнят определенные положения. Вы помните, что лоренцово преобразование имело вид

$$x' = \frac{x - vt}{\sqrt{1-\beta^2}}, \quad y' = y, \qquad z' = z, \qquad t' = \frac{t - \frac{vx}{c^2}}{\sqrt{1-\beta^2}}, \quad \left(\beta = \frac{v}{c}\right). \tag{21}$$

(Такой вид оно имело вследствие специального выбора координат).

Мы видели, что одно из самых существенных следствий из таких формул перехода от координат системы K к координатам системы K' заключается в том, что понятие времени и понятие длины потеряли абсолютное значение. Понятие времени и понятие длины уже не абсолютны для данного тела, а зависят от системы координат, в которой производятся измерения. Мы сначала установили все соотношения в каждой системе порознь, а затем высказали в качестве физического постулата, что если перенести твердый стержень из одной системы в другую, то он ведет себя согласно такого рода зависимости.

Одно следует отсюда сразу же: если длина стержня, перенесенного из одной системы в другую, изменяется в согласии с этой зависимостью, то, будучи перенесен из второй системы в первую, он опять будет иметь в первой системе те же самые размеры, как и до переноса. Иными словами, мы здесь постулируем ту не самоочевидную вещь, что длина стержня не зависит от истории его движения, т. е., будучи перенесен в другую систему, а потом обратно в первую, он не изменит своей длины.

Я хотел бы обратить ваше внимание еще на одно из следствий, вытекающих из наших формул. Если у нас имеется шар, находящийся в покоящейся системе, то этот шар, измеренный в движущейся системе, будет уже не шаром, потому что размеры в направлении движения изменились, а в поперечном направлении остались прежние. Это будет сплющенный эллипсоид. Отсюда следует, далее, что объем также является величиной, которая, будучи измерена в различных системах, имеет разное значение. Возьмите, например, параллелепипед; объем его равен произведению трех его измерений. Во время движения два измерения остаются без изменения, а одно сокращается; следовательно, объем изменится, и легко видеть, что мы будем иметь

$$V_{K'}^{K} = V_{K'}^{K'} \sqrt{1-\beta^2}.$$

Другими словами, объем тела также не есть инвариант, а зависит от системы отсчета.

Но у нас есть все же пока одно выражение, которое остается неизменным, будем ли мы его измерять в одной системе или в другой. Это не пространственный отрезок и не отрезок времени, а их комбинация: $x^2 - c^2t^2 = x'^2 - c^2t'^2$. Это выражение, составленное определенным образом из временных и пространственных координат, не изменяется, не зависит от системы координат, в которой вы измеряете.

Я хотел бы указать, что написанная форма инварианта обладает очень малой общностью, она специфична ввиду того, что мы

исходили из определенных систем координат. То, что мы исходили из определенной ориентации осей координат, несущественно, а существенно то, что мы подогнали начальную точку так, что при $x = 0$, $t = 0$ получается и $x' = 0$, $t' = 0$. Мы могли бы вести отсчет x' и t' от другого начала. Мы получим более общие выражения, если прибавим к x' некоторую величину α, а к t' — величину τ. Так как теперь, при $x = 0$, $t = 0$, x' и t' будут иметь какие-то произвольные значения, то при переносе начала координат выражение $x^2 - c^2t^2$ не останется инвариантным. Но теперь инвариантным будет выражение

$$(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - c^2 (\Delta t)^2 = \text{const},$$

где $\Delta x = x_2 - x_1$, $\Delta t = t_2 - t_1$ и т. д. Это очень существенно отметить, и вот почему. Инвариант является некоторым выражением, относящимся к *двум* событиям. Раньше могло показаться, что он относится к одному событию, здесь же сразу видно, что к двум. Это затемнялось только потому, что одно из этих событий было таково: $x = 0$, $t = 0$, и благодаря специальному выбору координат это переносилось и в другую систему. Теперь мы имеем более общий случай.

О скоростях мы больше говорить не будем. Мы видели, что формулы (21) предрешают вопрос о том, как преобразуются скорости, и видели, что теорема сложения скоростей приобретает новую форму. Я уже подчеркивал, что это получается потому, что мы даем определенное толкование сложению скоростей. Когда речь идет о сложении скоростей в одной и той же системе, то все остается по-старому, но при интересовавшей нас постановке вопроса обычная теорема сложения скоростей заменяется эйнштейновской, о которой мы достаточно говорили.

Обратимся теперь к довольно тонкому вопросу, который во всем дальнейшем играет существенную роль. Я ставлю вопрос следующим образом. Произошли два события, например две световые вспышки, причем в двух различных местах и в различные моменты времени. Мы знаем, что

$$x_2' - x_1' = \frac{x_2 + x_1 - v(t_2 - t_1)}{\sqrt{1-\beta^2}}, \qquad t_2' - t_1' = \frac{t_2 - t_1 - \frac{v}{c^2}(x_2 - x_1)}{\sqrt{1-\beta^2}}.$$

Спрашивается, могу ли я подобрать такую галилееву систему, движущуюся по отношению к исходной системе K, чтобы в ней эти события произошли в одной и той же точке, т. е. были бы «одноместными»? Мы знаем, что расстояния зависят от системы отсчета. Можно ли сделать два, вообще говоря, неодноместных (в данной системе K) события одноместными (в другой системе K')? Это

можно сделать, но не всегда. Для того чтобы это имело место, необходимо $x_2' - x_1' = 0$. Могу ли я подобрать скорость системы K' так, чтобы это было справедливо? Для этого нужно, чтобы числитель в выражении для $x_2' - x_1'$ равнялся нулю, т. е. $v = \frac{x_2 - x_1}{t_2 - t_1}$, где $x_2 - x_1$ — разность координат, а $t_2 - t_1$ — разность времен. Мы знаем, что система K' не может иметь скорость v, большую c. Это основное положение, ибо в противном случае мы получили бы мнимые величины. Значит, нашему требованию можно удовлетворить тогда и только тогда, если скорость $\frac{x_2 - x_1}{t_2 - t_1}$ меньше c. Обозначим $x_2 - x_1 = R$ и $t_2 - t_1 = T$. Следовательно, если для рассматриваемых событий $R < cT$ в какой-нибудь одной системе, то можно подобрать такую систему, в которой они будут одноместными. Если, наоборот, события подобраны так, что в какой-нибудь системе $R > cT$, то такой движущейся системы, в которой эти события были бы одноместными, нет. Другими словами, если два события таковы, что для них в какой-то системе $R > cT$, то это соотношение сохраняется во всех системах. Если же $R < cT$, то пространственное расположение неопределенно, в одной системе может быть $x_2 - x_1 > 0$, в другой может быть $x_2 = x_1$, а если скорость v будет еще больше, то может быть и так, что то, что произошло в одной системе справа, в другой произойдет слева.

Теперь поставим другой вопрос. Можно ли подобрать систему K' так, чтобы два события, не одновременные в данной системе K, в новой системе сделались одновременными? Мы требуем теперь $t_2' - t_1' = 0$, т. е. $\frac{x_2 - x_1}{t_2 - t_1} = \frac{R}{T} = \frac{c^2}{v}$. Очевидно, это возможно только тогда, когда $R > cT$.

Вы видите, что эти два случая исключают друг друга. Следовательно, все пары событий в смысле их пространственно-временных соотношений принадлежат к одному из двух классов.

Один класс — тот, где вопрос о том, имеет ли место $x_2 - x_1 > 0$ или $x_2 - x_1 < 0$,— чисто случайная вешь, и ответ зависит от системы отсчета, в которой события рассматриваются. Но зато, наверное, если в какой-либо системе K одно событие наступает позже, чем другое, то и во всякой другой системе оно случается позже.

Другой класс состоит из пар событий, для которых имеет место обратное: вы не можете сказать, случилось ли данное событие позже или раньше, чем другое,— это зависит от системы отсчета, но зато вы, наверное, можете сказать, что для этого класса пар $x_2 - x_1$ в любой системе отсчета либо положительно, либо отрицательно.

Таким образом, эти два класса содержат разнородные пары событий. Дальше мы увидим, какое это имеет значение.

Обратим теперь внимание на следующее: c — вполне определенная величина (и в этом весь смысл предыдущего рассмотрения), это скорость света в пустом пространстве. Представим себе, что c увеличивается до бесконечности, и посмотрим, во что обращается тогда лоренцово преобразование. Если $c \to \infty$, то (21) в пределе обращается в такую зависимость:

$$x' = x - vt, \quad y' = y, \ z' = z, \quad t' = t.$$

Это не что иное, как галилеево преобразование. Вы помните, как складывались скорости по Эйнштейну. Если $c \to \infty$, то $u' = u + v$, т. е. получается обычная теорема сложения скоростей. Таким образом, если рассматривать лоренцово преобразование и стремить c в бесконечность, то получается обычное галилеево преобразование — предельный, вырожденный случай лоренцова преобразования, которое представляет собой с математической стороны гораздо более общий случай. С физической стороны мы придаем ему значение именно потому, что считаем, что не существует скорости большей, чем c. Если бы существовала бесконечная скорость, то при тех же самых соображениях мы пришли бы к прежним галилеевым формулам. Вот почему говорят, что если бы в природе имелась бесконечная скорость сигнала, то все было бы «в порядке», преобразование было бы галилеевым и время во всех системах — одним и тем же.

Как же обстоит дело с возможностью сделать два события одновременными или одноместными в классике? Галилеево преобразование получается из лоренцова при $c \to \infty$. Условию $R < cT$ можно удовлетворить при этом всегда. Второму же условию $R > cT$ нельзя удовлетворить никогда. Значит, если бы мы пользовались галилеевым преобразованием, то мы всегда могли бы сделать два события одноместными, но никогда не одновременными. Это естественно. Сделать одновременными два в какой-либо системе разновременных события мы не можем потому, что одновременность в классике абсолютна. Ясно далее, почему любые два неодноместные события можно сделать одноместными. В классике это значит следующее: пусть происходит взрыв в одной точке и несколько позднее в другой. Вы едете в вагоне. Вы всегда можете сделать скорость такой, чтобы оба взрыва произошли против окна вагона, т. е. события будут одноместными. Это всегда можно сделать потому, что скорости не ограниченны. Какая бы ни была разница между временами, вы всегда можете взять скорость такой, чтобы в вашей системе отсчета события сделались одноместными. Из этого видно, что такого разделения событий на два класса в классике нет. В теории относительности оба класса полны, в клас-

сике же в одном классе находятся все объекты, а в другом вообще ничего нет.

Тем, что мы до сих пор рассматривали, физическая сущность интересовавших нас вопросов в той мере, в какой речь идет о кинематике, более или менее исчерпана. Но если идти дальше, если перейти к построению механики, электродинамики и т. д., то на весь вопрос о пространственно-временных отношениях интересно посмотреть с более общей точки зрения, по крайней мере формально более общей. Лично я не считаю возможным разделять в очень многих случаях формальную сторону и существо дела и поэтому в том, что дал Минковский, когда он, как говорят, сделал мир четырехмерным, вижу не только формальную сторону. Можно считать, однако, что это не больше чем способ вычисления, именно формальный способ, хотя сам Эйнштейн говорит, что без этого обобщения ему вряд ли удалось бы создать свой общий принцип относительности. Как бы то ни было, к этому вопросу нам нужно теперь обратиться.

Прежде всего несколько слов, хотя и не очень существенных, о более общем лоренцовом преобразовании, о котором до сих пор мы не говорили. Мы выбрали для сравнения очень специальные системы координат. У нас y и z оставались неизменными, а x был взят так, что начальные точки совпадали. Можно спросить себя, как преобразуются координаты и время, если взять две декартовы системы, ничем себя не связывая, т. е. взять и любое начало отсчета, и любое направление осей в каждой системе, и любое направление скорости одной системы по отношению к другой. Конечно, одно мы предполагаем по-прежнему,— что движение систем неускоренное, потому что мы рассматриваем только класс систем, движущихся друг по отношению к другу и по отношению к системе неподвижных звезд равномерно и прямолинейно. Как здесь будут выражаться координаты x', y', z', t' через x, y, z и t? Как здесь поставить вопрос, как искать это преобразование?

Можно и здесь по-прежнему исходить из тех же физических соображений, из которых мы исходили. Я напомню, в чем они заключаются. Мы требовали, во-первых, чтобы скорость света во всех системах была одна и та же и не зависела от движения источника и, во-вторых, чтобы изменение масштабов и времени при переходе от одной системы к другой было обратимо. Формально это означает, что если в одной системе мы имеем

$$(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - c^2 (\Delta t)^2 = 0,$$

т. е. распространение света по шару, то x', y', z', t' должны так зависеть от координат x, y, z, t, чтобы равенство этого выражения нулю в первой системе обращало бы его в нуль и во второй. Но мы видели, кроме того, что обратимость принципа относительно-

сти требует, чтобы не только равенство нулю этого выражения в одной системе влекло равенство нулю такого же выражения в преобразованных координатах, но чтобы это выражение было инвариантным, т. е.

$$(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - c^2 (\Delta t)^2 = (\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2 - c^2 (\Delta t')^2.$$

Мы ставим, следовательно, вопрос о том, чтобы найти соответствующие функции $x' = f_1(x, y, z, t)$, $y' = f_2(x, y, z, t)$ и т. д. Прежде всего они, как и раньше, должны быть линейны, потому что мы требуем, чтобы движение, равномерное и прямолинейное в одной системе, было таким же и в другой. Кроме того, формулы преобразования должны оставлять инвариантным наше квадратичное выражение. Этим уже почти полностью определены коэффициенты искомого преобразования, но здесь нужно обратить внимание на одну тонкость, относительно которой я хочу сказать теперь несколько слов и которая имелась уже и в нашем простом преобразовании. Мы имели

$$x' = \frac{x - vt}{\sqrt{1-\beta^2}}, \qquad t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1-\beta^2}}.$$

Я уже указывал, что если бы мы написали $x' = -\frac{x - vt}{\sqrt{1-\beta^2}}$, то требованию инвариантности наше выражение также удовлетворяло бы. Можно было бы взять со знаком минус и x' и t', наконец, можно было бы в первом случае взять плюс, а во втором минус, т. е. всего возможно четыре случая. Значит, если мы определим преобразование при помощи требования, чтобы оно оставляло инвариантным выражение $(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - c^2(\Delta t)^2$, то мы должны еще отдать себе отчет в том, почему мы берем x', y', z' и t' с определенными знаками. Я уже упоминал об этом, но сейчас это очень важно. Знаки могли бы быть другими. Чем мы руководствовались при их выборе?

Здесь есть один существенный момент и один менее существенный. Сначала — менее существенный. Предположим, что я написал бы в x' минус, и предположим, что я перехожу от одной системы к другой, у которой v становится все меньше и меньше и, наконец, $v = 0$. Я получим бы тогда $x' = -x$, а не $x' = x$. Это значит, что за положительное направление x в системе K' я выбрал другое направление, чем в системе K. Таким образом, если я выбираю положительный знак, то это означает не что иное, как условие, что при нашем специальном подборе систем координат прямая x' не только совпадает с прямой x, но и направление у них одинаково.

Когда мы берем более общий случай — преобразование всех трех координат, причем оси координат не имеют этого специального расположения, а произвольно повернуты, то мы можем поставить аналогичное требование, которое будет теперь заключаться в следующем: мы должны выбрать координатные системы так, что если одна из них является правой, то и все остальные должны быть правыми. Это чисто условное ограничение, от которого в дальнейшем придется отказаться.

Более существенный вопрос — выбор знака в t'. Если я поставлю минус, то при $v = 0$ я получу $t' = -t$. Это означает, что я обернул направление времени, т. е. в одной из систем, находящихся в покое друг по отношению к другу, t будет увеличиваться, а в другой уменьшаться; событие, которое в одной системе происходит раньше другого, в другой системе будет происходить позже, и наоборот. Это противоречит тому, чего требует причинность. Нужно, чтобы соотношения «позже» и «раньше» для причинно связанных событий сохранялись. Это требование не включено в распространение света, которое совместимо с тем, что преобразования могут изменять направление времени. Во всей метрике, во всей нашей постановке вопроса требования специального направления течения времени не содержится. Это всегда добавочное требование, каким, впрочем, оно является и во всей классической электродинамике. Поэтому здесь ничего нового нет, но на это нужно обратить внимание. Мы должны, таким образом, требовать также и того, чтобы при наших преобразованиях не могло быть «перевернутого» времени. А тогда, если поставить оба указанных условия, искомое преобразование определено уже вполне однозначно.

Введем новые обозначения. Нам теперь неудобно будет писать x', y', z', t'. Раньше я писал так потому, что в наших рассуждениях выделялось понятие времени, теперь же мы введем следующие обозначения:

$$x = x_1,\ y = x_2,\ z = x_3,\ ct = x_4.$$

Тогда искомое линейное преобразование можно написать следующим образом:

$$x_i' = \sum_k a_{ik} x_k + b_i, \qquad i = 1, 2, 3, 4. \tag{22}$$

Это запись самого общего линейного преобразования. Как выделить из него лоренцово преобразование? Оно выделяется требованием такого подбора a_{ik} и b_i, чтобы (22) оставляло инвариантным то квадратичное выражение, которое мы писали.

Что можно сказать относительно b_i — постоянных величин, которые присоединяются к однородным линейным функциям?

Требование инвариантности никакого ограничения на b_i не накладывает, они могут быть любыми числами. Действительно, ведь мы требуем инвариантности для выражения, составленного из разностей координат, а при вычитании величины b_i выпадают.

Для того, чтобы найти значения a_{ik}, нужно образовать выражение

$$(\Delta x_1')^2 + (\Delta x_2')^2 + (\Delta x_3')^2 - (\Delta x_4')^2,$$

подставить в него значения (22) и посмотреть, каким условиям должны удовлетворять a_{ik}, чтобы эта величина оставалась инвариантной. Легко показать, что условия будут такого же типа, как и в том случае, когда речь идет о преобразовании обычных декартовых координат в аналитической геометрии. Только там инвариантным является выражение $(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$, а у нас должно быть инвариантным выражение из четырех величин, причем четвертая берется с отрицательным знаком,— в этом заключается разница, но тип проблемы тот же самый. Если все это проделать, то получится, например, соотношение $a_{11}^2 + a_{21}^2 + a_{31}^2 - a_{41}^2 = 1$ и затем еще три подобных равенства. Далее мы найдем шесть уравнений вида $a_{11}a_{12} + a_{21}a_{22} + a_{31}a_{32} - a_{41}a_{42} = 0$.

Можно было бы написать условия для a_{ik} обычным образом со знаком суммы, но здесь мешает одно обстоятельство: как выделить последний член, имеющий знак минус? Мне кажется, существует очень простая запись. Я не могу записать инвариант в виде $\sum\limits_i (\Delta x_i)^2 = \text{const}$, потому что четвертый член входит с минусом. Будем обозначать это так:

$$\sum^{(4)} (\Delta x_i)^2 = \text{const} \qquad (23)$$

и будем считать, что там, где четверка встречается один или два раза, надо брать минус, а там, где она встречается три или четыре раза, то плюс. Если согласиться так писать, то условия для a_{ik}, которые необходимы, чтобы (22) выражало лоренцово преобразование, напишутся так:

$$\sum_k^{(4)} a_{ki}a_{kj} = \delta_{ij}, \qquad \text{где} \begin{cases} \delta_{ij} = 0 \text{ при } i \neq j, \\ \delta_{ij} = 1 \text{ при } i = j. \end{cases} \qquad (24)$$

Итак, для того чтобы преобразование (22) оставляло инвариантной форму (23), между a_{ik} должны быть 10 соотношений (24).

Большего, конечно, сказать нельзя; какие вы имеете каждый раз коэффициенты, об этом мы говорить не будем, но мы знаем теперь, как подойти к этому вопросу. Значит, формально дело обстоит так: лоренцово преобразование является линейным преобразованием, коэффициенты которого удовлетворяют определенным десяти условиям. Сколько вообще коэффициентов в преобразовании? Легко видеть, что в нем 20 коэффициентов. Так как они удовлетворяют десяти независимым уравнениям, то остается 10 степеней свободы. Таким образом, общее лоренцово преобразование есть линейное преобразование с 10 произвольными параметрами, в то время как в нашем лоренцовом преобразовании прежнего специального вида был всего один произвольный параметр — скорость v. Можно подсчитать на пальцах, что действительно должно быть 10 степеней свободы. Если мы имеем систему K и по отношению к ней движется система K', то сколько нужно степеней свободы, чтобы описать все возможные движения? Во-первых, начало координат произвольно, и это дает три параметра. Затем начало отсчета времени произвольно — это еще одна степень свободы; затем поворот осей определяется тремя углами и, наконец, имеются три компоненты относительной скорости, т. е. всего 10 параметров.

Это преобразование называется псевдоортогональным. Почему псевдоортогональным? Потому что тот же самый вопрос ставится в аналитической геометрии, где речь идет просто о повороте координатной системы и преобразование должно быть таким, чтобы оно сохраняло инвариантным расстояние, т. е. сумму трех квадратов разностей координат. Такое преобразование называется ортогональным. Здесь преобразование похоже на то, но сохраняет значение не суммы, а трех членов с положительными знаками и одного с отрицательным.

Итак, мы имеем многообразие ∞^{10} преобразований. Любое из них всегда можно получить путем такой последовательности операций. Для каких-либо систем, движущихся как угодно, вы можете сначала сделать самое обычное декартово преобразование координат, такое, чтобы оси систем сделались параллельны, т. е. вы можете сделать так, чтобы получилось преобразование только между одной координатой и временем. Выполнив его, вы можете еще раз повернуть координатную систему, придав ее осям первоначальное направление, т. е. вы делаете три последовательных операции. Если к любым системам применить последовательно эти три преобразования, то можно прийти к любому заданному положению. Все лоренцовы преобразования исчерпываются этими тремя последовательно примененными преобразованиями. Первое и последнее — это просто повороты координатных осей. Единственное, что ново,— это второй этап, т. е. специальное ло-

ренцово преобразование, которое мы уже изучили, а, значит, изучая его, мы косвенно изучили и общий случай, так что на нем не стоит более останавливаться.

Еще одно замечание. Требование инвариантности определяет a_{ik} не совсем однозначно, как и в случае специального лоренцова преобразования. Мы поставили требования, чтобы правая система оставалась правой и время сохраняло свое направление. Для первого нужно, чтобы детерминант (a_{ik}) был равен 1, а не -1. Для второго же необходимо, чтобы выполнялось условие $a_{44} > 0$. С этими двумя дополнительными условиями лоренцово преобразование полностью определено в самом общем случае. Но это скорее формальная вещь, к которой я все же хотел привлечь ваше внимание.

Теперь наступило время обратиться к несколько другому воззрению на все эти пространственно-временны́е соотношения. То, что вносит теория относительности, меньше, чем обычно думают, но все-таки нужно поговорить об этом подробнее.

Если независимо от того, стоим ли мы на классической точке зрения или на позициях теории относительности, всмотреться в то, что служит для измерения пространственно-временны́х отношений, что является элементом, из которого мы все складываем, элементом, при помощи которого мы характеризуем эти соотношения, то становится ясным, что это не точка пространства сама по себе и не момент времени сам по себе, а точка пространства в определенный момент времени. Как сказал как-то Минковский, никто никогда не видел точки в неопределенное время и не видел времени в неопределенной точке. Под последним я не подписался бы целиком, но, поскольку речь идет о физике, это, конечно, верно. Когда вы что-нибудь изучаете в физике, всегда дело касается какого-то момента времени в какой-то точке пространства. Поэтому элементом является не точка пространства и не время сами по себе, а *событие*, т. е. нечто такое, что определено *четырьмя* величинами — тремя пространственными координатами и моментом времени.

Если мы согласимся называть все эти четыре величины, которыми определено событие, координатами, то можно сказать, что элементами, из которых складываются наши пространственно-временные соотношения, являются четыре координаты. Совокупность четырех координат x, y, z, t Минковский называет мировой точкой. Это сделано, конечно, полностью по образу и подобию того, что делается в аналитической геометрии. В аналитической геометрии точкой называется совокупность трех чисел x, y, z. Вы, конечно, все это знаете, но я просто напомню то течение геометрической мысли, которое было наиболее полно оформлено в декартовой аналитической геометрии. Это так называемая арифметизация геометрии.

Тройка чисел называется в этой геометрии точкой. Дело рассматривается не так, что данная точка *имеет* координаты, а говорят, что сама точка — это чисто арифметическое, алгебраическое понятие. Пара троек чисел называется двумя точками, парой точек. Если рассматриваются точки (т. е. тройки чисел), удовлетворяющие алгебраическому уравнению $ax + by + cz + d = 0$, где a, b, c, d — постоянные, то говорят, что эти тройки чисел образуют плоскость. Опять-таки здесь нет никаких пространственных представлений. Просто я беру лишь те числовые тройки, которые удовлетворяют этому уравнению. Необходимо указать, что́ в этой геометрии называется расстоянием между двумя точками. По определению, расстоянием между двумя точками называется

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}. \qquad (25)$$

Так идет развитие аналитической геометрии, так можно построить замкнутую аналитическую систему, которая ставит себе, например, следующие задачи. Пусть даны две точки, т. е. даны две тройки чисел. Можно спросить: каковы те линейные преобразования этих чисел в другие две тройки чисел, которые оставляют инвариантным выражение (25)? Мы находим ортогональные преобразования, которые в аналитической геометрии мы называем преобразованиями координат.

Если так посмотреть на дело, то, в сущности, вся геометрия сводится к тому, что изучают совокупность таких преобразований, которые оставляют инвариантной величину (25). Можно доказать тогда, например, такие теоремы: даны два треугольника, т. е. заданы один раз три тройки чисел и другой раз три тройки чисел. Здесь опять нет ничего пространственного, треугольником называется совокупность трех троек чисел. Когда мы будем называть эти треугольники равными? Равными или конгруэнтными мы будем их называть тогда, когда при помощи одного преобразования, которое оставляет расстояния инвариантными, мы можем совместить первые три тройки со вторыми, т. е., применив это преобразование, получить из первых троек чисел вторые. И вот можно доказать, например, что если у вас есть дважды по три тройки чисел, которые в указанном смысле конгруэнтны, то вы можете определить понятие площади треугольников и доказать только из этих арифметических соображений, что их площади будут равны.

Такова одна из систем построения геометрии. Ничего общего с пространственными представлениями она пока что не имеет. В чем ее ценность? В том, что если вы возьмете какие-либо физические тела, например твердые тела, из этих твердых тел построите то, что мы называем треугольником, т. е. скрепите три стержня и возьмете координатную систему из трех стержней, то, найдя по

определенным рецептам координаты точек скрепления, т. е. получив для каждой точки тройку чисел, вы убедитесь, что эти тройки чисел, взятые из реальных предметов, будут подчиняться тем законам, которые вы вывели для троек чисел чисто абстрактно. Иными словами, утверждается, что если вы считаете перенос начала и поворот осей таким преобразованием, которое оставляет инвариантным (25), то такая геометрия применима к реальным предметам. Эйнштейн в своей общей теории относительности говорит, что это вовсе не так. Он построил *другую* систему алгебраических отношений, которую он назвал геометрией, и утверждает, что в природе применимы те соотношения, которые вытекают из нее. Вы видите, что при таком построении получается следующее: мы строим известную абстрактную систему, известные образы, понимаемые алгебраически, совершаем с ними разные операции и обнаруживаем, что этим операциям мы находим в реальной природе определенное соответствие. Именно поэтому такая система приобретает для нас ценность.

Из сказанного ясно, как получилось, что мы привыкли называть точкой не то, что мы представляем себе в пространстве, а просто три числа. Естественно и обратное: если мы видим, что в других соотношениях физики нам важны события, которые определяются четырьмя величинами, то, обобщая это понятие, мы называем точкой совокупность этих четырех чисел. Наша аналитическая геометрия определялась тройкой чисел. Соответственно здесь мы должны сказать, что соотношения, которые мы будем исследовать, отвечают пространству четырех измерений.

Я поставил вопрос в такой форме специально для того, чтобы показать вам следующее. Только привычка (для меня это несомненно) мыслить настоящую геометрию аналитически повела к тому, что в теории относительности, которая разбирает вопросы чисто аналитически, мы называем определенным образом полученные четыре числа x, y, z, t таким, я бы сказал, на первый взгляд невыгодным словом, как точка. Почему эти четыре числа называются точкой? Потому что алгебраизация обычной геометрии приучила нас к тому, что элемент из трех чисел мы называем точкой. Перенося этот способ дальше, мы говорим, что и в теории относительности мы станем называть систему чисел точкой. Но так как здесь четыре координаты, то мы строим *четырехмерный* мир. *Ничего большего это не значит.* Если мы говорим, что теория относительности рассуждает о четырехмерном мире, то ничего другого это не означает. *Ничего мистического, ничего метафизического в этой четырехмерности мира теории относительности нет.* Никто никогда не требовал, чтобы перестроились все пространственные представления, что если раньше был трехмерный мир, то теперь нужно научиться представлять себе четырехмерный.

Просто, когда встречаются такого рода задачи, где элементом служит нечто, определяемое тремя числами, это называется точкой, а когда у нас четыре числа, то мы также называем их точкой, но точкой четырехмерного пространства.

Я уже сказал, что элементом физического исследования является не точка пространства отдельно и время отдельно, а их совокупность. Такое понимание вовсе не принадлежит теории относительности. Это так же хорошо знали до всякой теории относительности. Эйнштейн это подчеркнул, но ничего нового здесь нет. И если хоть раз это понять и серьезно над этим подумать, то мир делается в указанном смысле четырехмерным, т. е. элементом оказывается нечто, определяемое четырьмя числами. Если вы говорите, что теория относительности утверждает четырехмерность мира, то я возражаю, что классика говорила то же самое. Новое заключается в другом. Если вы переходите от одной системы координат к другой, то классика считала, что из этих четырех чисел, которые определяют событие, три числа x, y, z преобразуются, т. е. изменяются отдельно, а время t остается неизменным. *Только в способе, каким преобразуются эти четыре числа* при переходе от одной системы отсчета к другой, заключается отличие между теорией относительности, с одной стороны, и классической физикой — с другой, а вовсе не в том, что мир стал четырехмерным, не в том, что физическими элементами, которые мы изучаем, являются события. Важно то, что при переходе от одной системы координат к другой t не остается неизменным, а x, y, z, t преобразуются совместно.

Результатом этого является следующее. Если в аналитической геометрии координаты x, y, z как таковые, сами по себе не являлись чем-то характеризующим данную пространственную точку, потому что они изменяются в зависимости от координатной системы, но зато расстояние в пространстве между двумя точками, отрезок, было чем-то абсолютным, что не изменялось, равно как и промежуток времени был также абсолютным, то теория относительности утверждает, что не только x, y, z являются, как и в геометрии, относительными, но и отрезок между двумя точками пространства также является относительным. Там он являлся абсолютным, здесь он относителен; промежуток времени являлся там абсолютным, здесь он тоже относителен. Не в том дело, что раньше все считалось абсолютным и только теперь поняли, что оно не абсолютно, а в зависимости от того, как на это смотреть, будет иметь то или другое значение. Нет, в прежней физике также знали, что есть и абсолютное и относительное. Относительными были координаты, абсолютным было расстояние между двумя точками и промежуток времени. Принцип относительности низвел и эти два понятия до уровня относительных понятий. Это мы знаем,

это нужно подчеркнуть, но если мы это подчеркиваем, то не надо забывать и гораздо более важную вещь,— что относительные понятия нужны нам, как леса для постройки. Это знали и раньше, но знали, кроме того, что есть нечто абсолютное, например расстояние и время. Они интересовали нас и с геометрической и с физической точки зрения. Интересовали потому, что это была абсолютная физическая реальность, это были инварианты, которые не зависят от системы координат.

И вот, благодаря тому что в классике были определенные инварианты, которые в теории относительности оказались не инвариантными, упор был сделан именно на эту сторону, на то, чтобы сказать: «Ага, а вот и то и это не инвариантно, все относительно и т. д.». Это создает ложную перспективу. Упор нужно сделать как раз на обратное, на то, что и теория относительности признает инварианты, что и с ее точки зрения существуют абсолютные вещи. Такой абсолютной вещью является выражение $(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - c^2(\Delta t)^2$, которое остается постоянным независимо от системы координат.

Таким образом, мы видим, что и в этом отношении принципиальной разницы нет. Теория относительности отнюдь не считает все относительным, точно так же раньше не все считалось абсолютным; она только перенесла центр тяжести в вопросе о том, что считать абсолютным, а что относительным, понимая под относительными те величины, которые меняют значение в зависимости от системы отсчета, а под абсолютными те, которые не изменяются при переходе от одной системы к другой. Теория относительности перенесла понятие абсолютного на новое выражение, которое получило специальное название — *интервал*. Это первое.

В современной литературе обычно называют интервалом корень квадратный из обратной величины, т. е. взятой со знаком минус

$$ds = \sqrt{c^2dt^2 - dx^2 - dy^2 - dz^2},$$

но нам удобнее пока писать по-прежнему

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2dt^2.$$

Интервал похож на инвариант аналитической геометрии (расстояние), но с той существенной разницей, что один знак здесь иной, чем все остальные. Это наводит нас на другой вопрос, о котором часто, особенно в популярных изложениях, говорится неправильно, что может привести к недоразумениям.

Вопрос заключается в следующем. Теория относительности — и в этом существенно новое, что она вносит,— при преобразова-

ниях от одной системы к другой преобразует и t и x, y, z *совместно*. Из этого делают такое заключение. Во-первых, мир четырехмерен. Мы уже знаем, что это ничего сверхъестественного не означает, что это только способ выражаться. Во-вторых, говорят, что четвертая координата есть время, и, таким образом, отождествляют время с пространственными координатами. Встречаешь иной раз такое выражение, что теория относительности, так сказать, уравняла четыре координаты: x, y, z, t. Оснований для этого я абсолютно не вижу. В том, что время — это одно, а пространство — другое, в том, что различие физической сущности этих понятий существует,— в этом отношении теория относительности ничего нового не дала. В некоторых книгах есть указания на разницу между временем и пространством с точки зрения психологии и теории познания. Конечно, так вопрос ставить можно и нужно, но я не понимаю только одного: почему его нужно ставить теперь и не надо было ставить раньше. Что, в сущности, здесь нового внесла теория относительности? То, что и t и x, y, z входят в преобразование совместно, могло, конечно, дать толчок в направлении к тому, чтобы подумать иначе, но никакого фактического основания для этого нет. Физик скажет просто так: рецепт для измерения расстояния один — приложение масштабов, отсчет числа, показывающего, сколько раз масштаб укладывается на измеряемом отрезке. Такой способ измерения дает нам то, что мы называем пространственными координатами. Способ измерения времени совершенно другой: вы берете периодический процесс (то, что называется часами) и при помощи его определяете время. Здесь совершенно различные приемы и процессы, и ни о каком равенстве или тождестве не может быть речи.

Сомнение в правильности этих разговоров о тождестве, особенно у таких людей, как я, закрадывается еще и по другой причине. Как я уже говорил, формальной стороне я придаю очень большое значение. Несомненно, в теории относительности так же как и в геометрии, все зиждется на инвариантах. Там нас интересуют в конечном счете не координаты какой нибудь точки, а расстояния, площади, углы между линиями. Это те быки, на которые все опирается, а не то, что преходяще и зависит от лесов. Точно так же и здесь центральным пунктом является интервал. Он не зависит от системы координат, и поэтому он нас больше всего интересует. И вот какое сомнение зарождается. В выражении интервала несомненно должна сказываться и сказывается структура всех понятий, которые в него входят. Не случайно в инвариантное расстояние аналитической геометрии все три координаты входят симметрично. Правда, могло бы быть, что они отражают различные понятия, но человеку, который придает формализму известное значение, который видит в формализме отображение

реальных вещей, должно бы показаться подозрительным, что при различии понятий имеется полная симметрия в формулах. Но этого как раз и нет. Все три координаты, по существу, тождественны. Здесь же, в случае интервала, структура не такая. Вы не можете сказать, что время — это то же самое, что и пространство, потому, мол, что оно входит симметрично с координатами. Нет, оно входит не симметрично, знак его иной. Это кажется не важным, ибо в математике не имеет значения, стоит ли плюс или минус, но это определяет все. Время входит не симметрично, и хотя бы поэтому сделать заключение о тождестве времени и пространства нельзя.

Наоборот, отсюда скорее следует, что они не тождественны. Но, как я уже сказал, я вообще не вижу, что нового внесла в этот вопрос теория относительности.

Итак, для нас существенно понятие интервала. Запишем интервал в виде

$$s^2 = c^2T^2 - R^2.$$

В этой четырехмерной геометрии (и это типично для теории относительности) в отличие от трехмерной один член со знаком минус. Если мы берем какие-нибудь два события (ведь интервал всегда характеризует соотношение между двумя точечными событиями), то для данной координатной системы мы получим какой-то квадрат пространственного расстояния и квадрат временнóго, которые в разных системах будут различны, но, какую бы систему мы ни взяли, интервал будет одинаковым. Предположим, что мы имеем в какой-нибудь системе $s^2 > 0$. Так как s^2 — инвариант, то для всех систем будет $s^2 > 0$. Другой раз может быть $s^2 < 0$, и тогда для всех систем $s^2 < 0$. Таким образом, все интервалы распадаются на два класса: квадраты интервалов могут быть абсолютно отрицательны и абсолютно положительны. Первое будет тогда, когда $cT > R$ по абсолютной величине, а второе — когда $cT < R$. Но ведь это как раз те условия, которые мы описали раньше, когда хотели сделать два события одноместными или одновременными. Значит, если два события таковы, что можно сделать их одноместными, то $R < cT$ и для этого случая квадрат интервала во всех системах будет положительным. Интервал, квадрат которого положителен, называется *времени-подобным интервалом*. Если же два события могут быть сделаны одновременными, то $R > cT$ и квадрат интервала отрицателен. Такой интервал называется *пространственно-подобным*.

Мы видели — и это необходимо еще раз подчеркнуть,— что наличие постоянного инварианта по меньшей мере так же важно, как и подчеркивание относительности того, что до теории отно-

сительности считалось абсолютным. Но если это так, то у физика сейчас же явится вопрос: существует ли вещественный представитель интервала?

Обратимся опять к геометрии. Возьмите две точки, например эти две лампы. Я нарочно беру лампы, так как между ними нет связи. Мы говорим, что расстояние между ними есть инвариант. Что мы считаем вещественным представителем этого расстояния? Можно сделать следующее: протянуть веревку или взять стержень достаточной величины, и вот этот стержень, на котором имеются две марки, зарубленные на местах обеих ламп, и есть физическое осуществление этого инварианта R.

Что же является физическим осуществлением инварианта в нашем случае? Это очень легко показать, но осуществление будет различно, смотря по тому, будет ли интервал времени-подобным или пространственно-подобным.

Возьмем сначала времени-подобный интервал. Я говорю так: возьмите какие-нибудь часы, покоящиеся в некоторой системе. Пусть теперь для простоты, чтобы все было совершенно наглядно и ничего нам не мешало, все происходит в темноте, а именно в какой-то момент происходит световая вспышка и затем вторая вспышка в той же точке. Пусть при первой вспышке мы отсчитали 12 часов, при второй, скажем, 12 ч. 5 м. Как измерить, чему равен интервал, соответствующий этим двум событиям? Конечно, я могу измерить интервал в различных системах. Но если я буду измерять в этой системе, где вспышки произошли в одном и том же месте, а значит расстояние равно нулю ($R = 0$), то я скажу, что для этой системы интервал есть просто $s = T$ (пусть $c = 1$). Если я буду измерять в другой, движущейся системе, то эти два показания неподвижных часов останутся инвариантом. Если я видел, что стрелки моих часов стоят при первой вспышке на 12, а при второй — на 12 ч. 5 м., то это факт, абсолютно инвариантный. Как бы ни двигались другие наблюдатели, которые проходят мимо *моих* часов, они всегда будут видеть на них один раз 12, а другой раз 12 ч. 5 м. Это и есть вещественный представитель интервала: разность показаний часов в системе, в которой место вспышки не движется. Итак, если я хочу найти физический процесс, показывающий величину инварианта s, то я беру часы, неподвижные в системе, в которой $R = 0$. В этой системе я измеряю s только этими часами, потому что в этой системе $R = 0$. Если я беру другую систему, то показания ее часов будут иные, но зато и R будет другим. Интервал же опять будет прежним. Значит вещественным представителем интервала является разность двух показаний часов, покоящихся в определенной системе, в которой события одноместны, или, как говорят, *собственное время* между этими событиями. Это вещественный представитель

интервала для случая $s^2 > 0$. Отсюда ясно, почему такой интервал называется времени-подобным.

Обратно, если два события заключаются в том, что в системе, в которой покоятся наши масштабы, одновременно на концах масштаба делаются две вспышки, то для таких двух событий в этой системе $T = 0$ и $s^2 = -R^2$. Поэтому длина масштаба, измеренная в неподвижной системе, определяет пространственно-подобный интервал. Таким образом, неподвижные часы измеряют времени-подобный интервал, а неподвижный масштаб измеряет пространственно-подобный интервал.

Итак, процессы, которые являются вещественными представителями этого инварианта, найдены. В общем случае как часы, так и масштабы измеряют интервал. И если права теория относительности, то она показывает, что понятие интервала вовсе не абстрактное понятие, что существует его вещественные представители — часы и масштабы, которые измеряют не время отдельно и пространство отдельно, как думали до сих пор, а интервал.

В этом отношении неудачно само название «теория относительности», так как оно заставляет думать, будто эта теория делает все относительным. Наоборот, она признает абсолютные величины, как признавали их и прежде, но она переносит акцент и делает относительными те величины, которые раньше считались абсолютными.

Я хотел бы, чтобы вы ясно видели, что в четырехмерном пространстве нет ничего, кроме способа выражения, что теория относительности не сказала ничего такого, откуда следовало бы, что время сделалось тождественным или почти тождественным пространству, и что новое заключается только в том, что передвинуты некоторые понятия, что указаны новые величины, которые теория относительности считает абсолютными.

ТРИНАДЦАТАЯ ЛЕКЦИЯ

(28.V 1934 г.)

Краткое резюме. Понятие группы преобразований. Матрица лоренцова преобразования. Умножение преобразований. Определение группы, примеры. Непрерывные группы. Лоренцовы преобразования образуют группу. Что дает требование групповости для установления вида преобразований

Элементом пространственно-временных соотношений является событие, т. е. чрезвычайно кратковременный процесс в очень малой области пространства, подобный вспышке или чему-нибудь в этом роде. Событие характеризуется четырьмя величинами — тремя пространственными величинами и временем, и можно сказать, что событие характеризуется четырьмя координатами. Это не ново, и классика тоже знала, что для определения события нужны четыре числа. Три из них входят во все соотношения симметрично и совершенно однородны, четвертое — время — совершенно иначе физически определено,— это величина совершенно другого характера. Теория относительности и в этом отношении ничего не меняет, для нее время также по существу отлично от пространственных координат.

Весь интерес был сосредоточен на вопросе о том, как нужно переходить от одной системы координат к другой. Тут сказывается существенное отличие ответа теории относительности от ответа, который давала классика.

Классика отвечала так: координата t остается неизменной для всех движущихся систем координат, и мы имеем право сказать, что время есть абсолютная координата; x', y', z' зависят от x, y, z, но благодаря тому, что t остается инвариантным, получается разделение на пространственные и временны́е координаты. Поэтому четырехмерность — то, что всегда необходимы четыре координаты,— стушевывалась, разбивалась на трехмерность и одномерность, т. е., грубо говоря, были не 4 координаты, а $3 + 1$.

Теория относительности говорит, что это неверно, что при переходе от системы K' к системе K преобразование координат происходит так, что не только $x' = f_1(x, y, z, t)$, $y' = f_2(x, y, z, t)$, $z' = f_3(x, y, z, t)$, но и $t' = f_4(x, y, z, t)$. Таким образом, разделение на время и пространство однозначным образом невозможно, так как зависит от системы, к которой вы переходите. Благодаря тому, что четыре координаты здесь сплетены, что их объединение это больше, чем декларация, что с этим приходится работать и к этому приспособить весь аппа-

рат, в специальной теории относительности говорят всегда о четырехмерном мире.

Исходя из определенных физических постулатов — относительности, с одной стороны, и независимости скорости света в вакууме от движения источника — с другой, — нам удалось найти вид этих функций. Формулы перехода от одной системы к другой выразились в форме лоренцова преобразования. Это преобразование линейно и в самом общем случае записывается так:

$$x_i' = \sum_k a_{ik} x_k + b_i,$$

причем для того, чтобы удовлетворялись два основных постулата, коэффициенты должны удовлетворять определенным соотношениям, а именно таким, чтобы преобразование было псевдоортогональным. Величины b_i произвольны и определяют только перенос начала координат и изменение отсчета времени. Мы всегда будем считать, что начало координат и отсчета времени в обеих системах совпадают и тогда $b_i = 0$, так что мы получаем преобразование

$$x_i' = \sum_k a_{ik} x_k,$$

где 16 коэффициентов удовлетворяют 10 определенным условиям, т. е. остается 6 произвольных параметров. Совокупность таких лоренцовых преобразований отображает возможность того, что новая система K' может быть повернута по отношению к системе K, что дает три угла поворота, например три эйлеровых угла, и может двигаться с произвольной постоянной скоростью, что дает три компоненты этой скорости. Эти величины и составляют 6 параметров. Мы видели, что всякий переход от одной системы координат к другой можно составить из трех типов преобразований: простого поворота осей, как в аналитической геометрии, затем специального лоренцова преобразования, затрагивающего одну координату и время, и, наконец, еще одного поворота осей.

Значит, то, что нас интересует,— это специальное лоренцово преобразование

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c_2}}}, \qquad y' = y, \qquad z' = z, \qquad t' = \frac{t - \frac{v}{c^2}\,x}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Именно оно специфично для теории относительности, остальное же складывается из этого преобразования и простых поворотов.

Отсюда мы сделали ряд выводов, кажущихся на первый взгляд парадоксальными, но в которых на самом деле ничего парадоксального нет. Во-первых, длина стержня не является чем-то абсолютным, а зависит от системы, в которой вы ее измеряете. Понятие времени также утрачивает абсолютность и зависит от системы, в которой вы измеряете. Существенным было также то, что теорема сложения скоростей в обычной форме несправедлива и заменяется эйнштейновской теоремой сложения скоростей.

Что будет занимать нас сегодня и о чем я хотел немного рассказать,— это некоторые замечательные свойства преобразований, рассматриваемых не каждое в отдельности, а как совокупность, т. е. некоторые свойства совокупности лоренцовых преобразований.

Для этого нам нужно будет рассмотреть предварительно несколько полуформальных вещей. Мы познакомимся с очень существенным понятием — с понятием *группы преобразований* и вообще с понятием *группы*.

Понятие группы в современной математике и физике играет все бо́льшую и бо́льшую роль. Мы можем коснуться его только самым поверхностным образом. В учебниках по теории относительности почти всегда говорится, что лоренцовы преобразования образуют группу, и на этом дело кончается. Мне кажется, что на этом нужно остановиться немного подробнее, тем более что после того, как Дирак ввел релятивистское уравнение в квантовой механике, специальная теория относительности несомненно получила известное расширение, а понятие группы позволяет охватить это расширение и понять, в каком смысле нужно пойти дальше и как далеко можно с таким расширением идти. Поэтому мне и хотелось остановиться на этом вопросе больше, чем это обычно делается.

Сначала несколько предварительных замечаний. Лоренцово преобразование — это линейное преобразование, причем мы будем рассматривать только однородное преобразование. Такое преобразование вполне охарактеризовано, если даны его коэффициенты a_{ik}. Квадратная таблица

$$\begin{pmatrix} a_{11}a_{12} & . & . & . & a_{1n} \\ . & . & . & . & . \\ . & . & . & . & . \\ a_{n1}a_{n2} & . & . & . & a_{nn} \end{pmatrix},$$

в которой каждый коэффициент стоит на своем месте, называется, как вы знаете, матрицей. Если задана матрица, то тем самым задано преобразование, и изучение преобразований сводится, таким образом, к изучению таких матриц. Мы хотим перейти теперь

к вопросу о свойствах совокупности, т. е. сравнивать различные матрицы между собой.

Представьте себе, что вы хотите перейти от системы K к системе K', движущейся с определенной скоростью по отношению к системе K. Мы будем записывать это таким образом. Матрицу, написанную выше, будем обозначать одной буквой A и символически запишем наше преобразование так:

$$x' = A(x)$$

(обычно скобок не пишут). Если матрица нам дана, то мы сразу можем написать преобразование. Вы видите, что чем дальше мы продвигаемся, тем проще и проще становится запись. Много десятков лет назад выписали бы все четыре уравнения полностью. Потом перешли к такой записи $x_i' = \sum a_{ik}x_k$, где, приписывая индексы в одном уравнении, охватывают все уравнения. Это уже гораздо более простая запись. Теперь это записывается совсем просто, но, конечно, по существу это то же самое.

Итак, мы сделали переход от координат системы K к координатам системы K'. Пусть имеется теперь третья система, которая движется по отношению к K' (а также по отношению к K),— система K''. Я хочу выразить события в K'' через координаты K'. Пусть скорость системы K'' по отношению к K' есть v_1. Тогда я могу написать $x'' = B\,(x')$, где B — новая матрица. Напишем матрицу A для наших специальных преобразований. Она такова:

$$\begin{pmatrix} \frac{1}{\sqrt{1-\beta^2}} & 0 & 0 & -\frac{v/c^2}{\sqrt{1-\beta^2}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{v/c^2}{\sqrt{1-\beta^2}} & 0 & 0 & \frac{1}{\sqrt{1-\beta^2}} \end{pmatrix}, \qquad \beta = \frac{v}{c}.$$

Если мы переходим от системы K' к системе K'', то все будет выглядеть так же, но с матрицей B, т. е. вместо v всюду будет v_1. Таким образом, сначала мы перешли от системы K к системе K', а затем от системы K' к системе K''. Но мы могли бы прямо перейти от K к K''. Как выразятся координаты K'' через кординаты K, если мы знаем, как выражаются координаты K' через координаты K и координаты K'' через координаты K'? У нас есть

$$x_i' = \sum_k a_{ik}x_k, \text{ или } x' = A\,(x)$$

и

$$x_i'' = \sum_k b_{ik} x_k', \quad \text{или} \quad x'' = B(x').$$

Мы можем подставить во второе преобразование x' из первого и получим опять линейные функции от x, но с другими коэффициентами. Весь вопрос заключается в том, как выражаются коэффициенты результирующего преобразования, составленного из двух первых преобразований, которое дает переход от системы K к системе K''. Я могу написать

$$x_i'' = \sum_k c_{ik} x_k, \quad \text{или} \quad x'' = C(x).$$

Спрашивается: как вычислить матрицу C, если даны A и B? Это простая арифметическая задача, и легко показать, что элемент матрицы C, скажем c_{ik}, составляется из соответствующих элементов a и b следующим образом:

$$c_{ik} = \sum_l b_{ij} a_{jk}. \tag{26}$$

Символически это записывается так:

$$C = BA,$$

и говорят, что C есть произведение двух первых матриц.

Физически третья матрица C означает такое преобразование координат, которое соответствует переходу от системы K к системе K'', выраженному при помощи двух переходов — от K к K' и от K' к K''.

Формула (26) вам хорошо знакома: если бы у вас здесь были не матрицы, а детерминанты

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \quad \text{и} \quad \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix},$$

то элементы детерминанта, являющегося произведением двух таких детерминантов, выражались бы именно таким образом, т. е. для произведения матриц мы имеем то же правило, что и для произведения детерминантов. Я хотел бы только подчеркнуть существенную разницу между детерминантами и матрицами. Детерминант — это число. Я записываю его в виде таблицы, но это одно число. Матрица — это таблица из n^2 чисел. Если я напишу вместо

первой таблицы такую

$$\begin{vmatrix} a_{22} & a_{12} \\ a_{21} & a_{11} \end{vmatrix},$$

то как детерминант она равна первой, а как матрица не равна, потому что для матрицы строго определен порядок элементов, здесь же порядок изменен. Об этой разнице никогда не следует забывать. Для всякой матрицы можно вычислить детерминант, составленный из ее элементов. Детерминант матрицы, являющейся произведением двух матриц, равен произведению детерминантов этих двух матриц. Это то же самое. Но вот где существенная разница — и это очень типично для матриц. Вообще говоря, если сначала вы делаете преобразование B, а затем A, то получите другую матрицу, чем в том случае, если сначала сделать преобразование A, а потом B. Для детерминантов это не играет роли, но в произведении матриц это будут различные вещи. Поэтому, называя новую матрицу произведением двух матриц, нужно быть осторожными, потому что это произведение не обладает коммутативностью. Но вы уже привыкли к тому, что не все произведения обладают коммутативностью; например, векторное произведение не коммутативно. Между прочим, вся квантовая механика основана на некоммутативных произведениях.

Итак, если нам нужно перейти от системы K к системе K'', то сначала мы можем перейти к системе K', а затем от K' к системе K''. Мы имеем вполне определенный рецепт, как найти преобразование от K к K'', когда оба промежуточных преобразования нам известны. Нужно последовательно и в определенном порядке применить оба этих преобразования, сначала A, потом B, и мы получим новое преобразование C. Мы научились, как по двум преобразованиям найти третье, которое из них составлено, т. е. является результатом их последовательного применения. Это третье преобразование мы называем произведением первых двух. Пока запомним это и обратимся к тому, что нас в данном случае интересует.

Совокупность всех лоренцовых преобразований обладает замечательным свойством, а именно является группой. Прежде всего, что такое группа? Позвольте остановиться на этом и дать, может быть, несколько поверхностное, недостаточно углубленное и даже не всегда достаточно строгое определение понятия группы в общем случае, а потом применить его к нашим вопросам.

Пусть у нас есть совокупность каких-нибудь объектов в конечном или бесконечном числе. Будем называть эти объекты *элементами*. Это могут быть самые разнообразные вещи — числа, матрицы, преобразования как линейные, так и нелинейные. Пусть установлено правило, как каждой паре этих элементов отнести

новый элемент, причем пары, вообще говоря, направлены, т. е. может быть так, что паре AB будет отнесен другой элемент, чем паре BA, но каждой паре, составленной по определенному закону, будет отнесен новый элемент. Возьмем самый простой пример. Пусть речь идет о совокупности целых положительных чисел 1, 2, 3, 4,... Условимся каждой паре чисел относить их сумму. Если я возьму числа 1 и 3, я буду относить им число 4, если я возьму 1 и 2, то отнесу им 3. Я установил правило, по которому каждой паре элементов отношу третий элемент. Я мог бы взять вместо сложения умножение. Тогда элементам 1 и 2 был бы отнесен элемент 2, элементам 2 и 3 — элемент 6 и т. д. Этот новый элемент, который я по установленному правилу отношу к любой паре, я буду называть *произведением* этих двух элементов. Это произведение не в том смысле, как мы понимаем обычно, так как я могу, например, установить правило сложением, т. е. буду относить третий элемент тем, что буду складывать пару элементов, но я все же буду говорить, что это произведение в новом смысле этого понятия. Я могу так условиться и так условилась определять вся теория групп.

Если у нас есть совокупность матриц, то матрицам A и B сопоставляется матрица C, определяемая по способу, который получился из физических соображений о том, какая матрица должна выражать последовательно произведенные преобразования A и B. Матрица C называется произведением матриц A и B, а закон ее составления и есть тот рецепт, который позволяет мне каждой паре элементов сопоставлять и здесь новый элемент.

Если совокупность элементов такова, что каким-либо образом определенное произведение любой пары этих элементов является опять элементом совокупности, то мы говорим, что эта совокупность образует группу. Это — главное свойство. Его недостаточно, но оно необходимо и является основным свойством группы. Кроме того, должны быть соблюдены еще следующие условия. Я знаю, как из A и B получить элемент AB — рецепт произведения,— и я знаю, что AB принадлежит к той же совокупности. Я могу этот новый элемент AB сопоставить с третьим элементом C: $(AB)\ C = D$. Но я мог бы сначала сопоставить B и C и затем образовать $A\ (BC) = D'$. Требуется, чтобы закон сопоставления оба раза давал одно и то же ($D = D'$), т. е. чтобы был справедлив *ассоциативный закон.*

Без этого требования тоже обойтись нельзя, оно очень существенно. Но, как сказано, самое существенное то, что произведение двух элементов есть опять элемент группы. Отсюда возникает само понятие группы — вы не выходите за пределы совокупности.

Далее требуется, чтобы в числе элементов совокупности был такой элемент, произведение которого на любой элемент A давало бы опять A. Такой элемент называется *единицей* группы или

единичным элементом. Мы имеем правило, по которому можем из любых двух элементов получать третий. Требуется, чтобы существовал такой элемент, который, будучи сопоставлен по установленному правилу с другим элементом, давал бы произведение, не отличающееся от этого второго элемента.

Наконец, последнее условие. Требуется, чтобы для каждого элемента A совокупности всегда можно было найти такой элемент B, что произведение AB равно единичному элементу. Тогда элемент B называется обратным A и обозначается A^{-1}. Это взято из обычной символики. Таким образом, если есть совокупность элементов, в которой установлено правило умножения, если произведение всегда принадлежит к элементам совокупности и если соблюдены указанные добавочные условия, то такую совокупность мы называем группой. Какие объекты взяты в качестве элементов, совершенно безразлично, важны только указанные свойства объектов.

Позвольте дать простые примеры для пояснения. Возьмем в качестве элементов числа 0, 1, 2, 3, 4, 5, 6, 7. В качестве умножения возьмем сложение этих элементов. Если я спрошу, какой элемент соответствует элементам 2 и 3, то это будет 5. У меня есть совокупность, в которой установлен закон получения произведения. Является ли эта совокупность группой? Нет, это не группа. Возьмем, например, элементы 4 и 5. Произведение 4 и 5, по определению, есть 9, но это число не входит в совокупность. А ведь главным нашим требованием было, чтобы всякое произведение также входило в совокупность. Возьмем тот же ряд чисел, но понятие умножения определим другим образом: мы будем складывать числа, как и раньше, но произведением будем считать не сумму, а остаток суммы от деления на 8. Если теперь я возьму два каких-нибудь числа из совокупности, например 5 и 6, то получу 11, а разделив на 8, получу в остатке 3. Это число входит в совокупность. Значит, с таким определением умножения наша совокупность является группой. Во всяком случае, первый признак группы здесь имеется.

Почему важны такие совокупности? Оказывается, что совокупности, обладающие свойствами группы, имеют ряд других чрезвычайно существенных свойств.

Проверим другие свойства нашей группы. Имеется ли в нашей совокупности единица? Единица определена тем, что, будучи сопоставлена с любым другим элементом, любым другим числом нашей совокупности, она дает то же число. Что здесь будет единицей? Очевидно, нуль, потому что, будучи прибавлен к другому числу, он не изменяет этого числа. Имеются ли здесь обратные элементы, которые, будучи сопоставлены с некоторыми элементами, дают нашу единицу, т. е. нуль? Обратные элементы здесь есть. Это числа, одинаково отстоящие от обоих концов строки (не считая нуля).

Так, например, для единицы обратным элементом будет 7, потому что, сопоставив их, мы получим нашу единицу, т. е. нуль. Обратным элементом для 2 будет 6. Таким образом, эта простая совокупность, при указанном определении произведения, является группой. Группы такого типа называются *конечными*, потому что они содержат конечное число элементов.

Могут быть и *бесконечные* группы с бесконечным числом элементов. Например, совокупность всех целых чисел, положительных и отрицательных, включая нуль, образует группу, если под произведением понимать алгебраическую сумму.

Нас будут интересовать *непрерывные* группы. Под непрерывной группой мы будем понимать следующее. Каждый элемент характеризуется каким-нибудь параметром, который может принимать в известной области всевозможные значения. Один элемент от другого будет отличаться тем, что параметр принимает различные значения, а так как параметр, по предположению, изменяется непрерывно, то мы получим непрерывный ряд элементов. Такими элементами являются и преобразования Лоренца. Что в них является параметром? Скорость. Для данной скорости у вас всегда имеется определенный элемент — определенное преобразование. Измените скорость, и у вас будет другое преобразование, соответствующее новому параметру. Так как скорость меняется непрерывно между $-c$ и $+c$, то получается непрерывная цепь элементов, которую мы и будем рассматривать как совокупность. Если эта совокупность удовлетворяет всем требованиям, которые мы ввели, т. е. если сначала установлено понятие умножения и затем оказывается, что совокупность удовлетворяет нашим требованиям, то такую непрерывную совокупность элементов мы назовем непрерывной группой.

Лоренцовы преобразования образуют группу. Я утверждаю, что если бы этого не было, то получились бы большие неприятности, получилось бы так, что мы не знали бы, с чего начать при физическом подходе. Хотя мы и не ставили перечисленных условий, хотя мы и не требовали группового характера от преобразований Лоренца, но мы увидим, что без этого вряд ли можно было бы построить какую бы то ни было последовательную систему.

Итак, непрерывными группами мы называем совокупность элементов, изменяющихся непрерывно и характеризующихся одним или многими параметрами, причем в такой совокупности выполняются перечисленные нами условия. Если элемент совокупности зависит от R параметров, каждый из которых может изменяться независимо от других, то мы говорим о R-членной группе. Например, лоренцовы преобразования самого общего вида имеют 10 независимых параметров и образуют десятичленную группу. Однородные преобразования образуют шестичленную группу,

а специальные лоренцовы преобразования составляют одночленную группу — с одним параметром v.

Итак, в качестве элемента совокупности мы будем рассматривать преобразования. Произведением мы назовем преобразование, получающееся в результате двух последовательно примененных, преобразований. Тогда основное свойство совокупности преобразований, которое делает ее группой, мы можем выразить несколько иначе: совокупность преобразований представляет собой группу, если последовательное проведение двух преобразований всегда может быть заменено одним преобразованием, также относящимся к совокупности. Кроме того, в совокупности должны быть налицо единичное (т. е. тождественное) преобразование и обратные преобразования.

По существу нас интересует лишь специальное лоренцово преобразование, которое оставляет y и z неизменными и преобразует только x и ct и в котором уже выявляется вся физическая сторона дела. Поэтому мы будем заниматься только этими одночленными преобразованиями, причем будем рассматривать преобразования только x и ct. Значит, матрица лоренцова преобразования такова:

$$\begin{pmatrix} \dfrac{1}{\sqrt{1-\beta^2}} & -\dfrac{\beta}{\sqrt{1-\beta^2}} \\ -\dfrac{\beta}{\sqrt{1-\beta^2}} & \dfrac{1}{\sqrt{1-\beta^2}} \end{pmatrix}.$$

Мы хотим показать, что это есть группа. Но сначала несколько слов о совокупности преобразований с непрерывно изменяющимся параметром.

Если вы возьмете любое преобразование с изменяющимся параметром, то, вообще говоря, совокупность таких преобразований не является группой; группа — это очень специальная совокупность, и в этом весь интерес. Я приведу простые примеры преобразования, которые являются группой, и таких, которые группой не являются. Возьмем преобразования

$$x' = x + k,$$
$$y' = y + 2k.$$

Это совокупность преобразований с изменяющимся параметром k, где k — любое вещественное число. При каждом фиксированном значении k это элемент совокупности — одно преобразование. Спрашивается, является ли эта совокупность группой? Имеется ли здесь единица? Очевидно, да. Это преобразование с $k = 0$. Но как обстоит дело с главным свойством, с тем, что последовательное применение двух преобразований опять является преобразованием,

относящимся к этой же совокупности? Это свойство также присутствует. В самом деле, возьмем два преобразования из нашей совокупности:

$$\left.\begin{aligned} x' &= x + a, \\ y' &= y + 2a, \end{aligned}\right\} \quad \left.\begin{aligned} x'' &= x' + b, \\ y'' &= y' + 2b \end{aligned}\right\}$$

и образуем произведение, т. е. сделаем эти два преобразования последовательно. Подставив x' и y' во вторые две формулы, получим

$$\left.\begin{aligned} x'' &= x + a + b, \\ y'' &= y + 2\,(a + b). \end{aligned}\right\}$$

Очевидно, это преобразование также принадлежит к нашей группе, причем параметром здесь является $c = a + b$, и, следовательно, наша совокупность преобразований есть группа. Если вы возьмете $b = -a$, то получите единичное преобразование, так что ко всякому преобразованию можно подобрать обратное ему. Это свойство группы тоже удовлетворено.

Когда берут такие элементарные примеры, то на первый взгляд кажется, что любые преобразования образуют группу. Легко убедиться в обратном. Возьмем преобразование

$$\begin{aligned} x' &= x + k, \\ y' &= y + k^2. \end{aligned}$$

Совокупность таких преобразований не является группой. Возьмем частные случаи $k = a$ и $k = b$, т. е. возьмем два преобразования:

$$\left.\begin{aligned} x' &= x + a, \\ y' &= y + a^2, \end{aligned}\right\} \quad \left.\begin{aligned} x'' &= x' + b, \\ y'' &= y' + b^2. \end{aligned}\right\}$$

Если теперь произвести подстановку, то получим

$$\left.\begin{aligned} x'' &= x + a + b, \\ y'' &= y + a^2 + b^2. \end{aligned}\right\}$$

Это преобразование не попадает в нашу совокупность, так как для этого во втором уравнении должен был бы стоять квадрат параметра, входящего в первое уравнение, а так его подобрать мы здесь не можем. Таким образом, любые преобразования с любой функцией. вообще говоря, не являются группой. Можно сформулировать в общем случае, когда совокупность преобразований является группой. Пусть у нас есть какое угодно нелинейное преобразование, которое мы запишем так:

$$x_i' = f_i\,(x_1,\ x_2, \ldots;\ a,\ b,\ c)$$

и пусть второе преобразование будет

$$x_i'' = f_i\,(x_1', \ x_2' \ldots; \ a', \ b', \ c').$$

Мы можем подставить сюда x' из первого преобразования и получим тогда некоторую функцию от x, a, b, c, a', b', c'

$$x_i'' = F_i\,(x_1, x_2, \ldots; \ a, \ b, \ c; \ a', \ b', \ c').$$

Если в результате этой подстановки x'' выразятся той же *самой* функцией f_i, но с другими параметрами, которые как-то зависят от a, b, c, a', b', c'

$$a'' = \psi_1\,(a, \ b, \ c; \ a', \ b', \ c'),$$
$$b'' = \psi_2\,(a, \ b, \ c; \ a', \ b', \ c'),$$
$$c'' = \psi_3\,(a, \ b, \ c; \ a', \ b', \ c'),$$

т. е. если совокупность двух подстановок можно заменить одной подстановкой из той же совокупности с другими значениями параметров, то мы скажем, что эти преобразования образуют группу.

Если лоренцовы преобразования образуют группу, то мы можем сказать, что группу образуют матрицы лоренцовых преобразований,— это то же самое, потому что в случае линейных преобразований матрица полностью характеризует преобразование. Мы хотим показать, что лоренцовы преобразования образуют группу в общем случае. Чем определены лоренцовы преобразования? По определению, это линейные преобразования, оставляющие инвариантной форму

$$(\Delta x_1)^2 + (\Delta x_2)^2 + (\Delta x_3)^2 - (\Delta x_4)^2.$$

Я произвожу первое преобразование, и при этом

$$(\Delta x_1)^2 + (\Delta x_2)^2 + (\Delta x_3)^2 - (\Delta x_4)^2 = \\ = (\Delta x_1')^2 + (\Delta x_2')^2 + (\Delta x_3')^2 - (\Delta x_1')^2.$$

Второе лоренцово преобразование характеризуется тем, что

$$(\Delta x_1')^2 + (\Delta x_2')^2 + (\Delta x_3')^2 - (\Delta x_4')^2 = (\Delta x_1'')^2 + (\Delta x_2'')^2 + (\Delta x_3'')^2 - (\Delta x_4'')^2.$$

Я произвожу оба преобразования последовательно, и так как они оба линейны, то в результате я опять получу линейное преобразование. Спрашивается, является ли результирующее преобразование также лоренцовым? Если да, то мы почти доказали, что это группа, потому что это основное свойство группы.

Но легко видеть, что из двух написанных равенств следует

$$(\Delta x_1)^2 + (\Delta x_2)^2 + (\Delta x_3)^2 - (\Delta x_4)^2 = \\ = (\Delta x_1'')^2 + (\Delta x_2'')^2 + (\Delta x_3'')^2 - (\Delta x_4'')^2,$$

т. е. совокупность преобразований также оставляет эту форму неизменной, а это и есть характеристика лоренцовых преобразований. Легко проверить, что и все остальные свойства группы здесь также соблюдены,— я не буду на этом останавливаться. Итак, мы можем сказать окончательно, что лоренцовы преобразования образуют группу.

Что было бы, если бы они не составляли группы? Лоренцовы преобразования позволяют находить координаты одной движущейся системы в функции от координат другой системы, по отношению к которой данная система движется. Представьте себе, что у нас имеются три координатные системы K, K' и K'' и мы хотим найти зависимость координат x'' от координат x. Мы можем сначала найти, как зависят x от x', потом — как зависят x' от x'', и, наконец, найдем зависимость x от x''. Это первый путь, совершенно законный. Но, с другой стороны, мы можем не учитывать, что имеется система K', она может нас не интересовать. Если есть только две движущиеся друг относительно друга системы K и K'', то принцип относительности дает возможность найти x'' сразу как функцию от x. Представьте себе теперь, что преобразования не образуют группы. Тогда по первому способу мы нашли бы зависимость x'' от x, которая не соответствовала бы тому, что мы получили во втором случае. Совокупность преобразований позволяет переходить от любого преобразования к любому. Если я к третьему преобразованию перешел при помощи промежуточного и получил бы преобразование, которое не относится к совокупности, т. е. не выражает перехода от одной движущейся системы к другой, то я нашел бы некоторую зависимость между координатами x'' и x, но это не была бы зависимость координат одной движущейся системы от координат другой, а ведь мы должны были бы получить такую зависимость. Значит, все наше построение было бы непоследовательным, нам запрещалось бы сравнивать сначала x и x', а затем x' и x''. Не было бы транзитивных соотношений, которые позволяют нам строить всю нашу концепцию, и эта концепция обладала бы провалом, который, быть может, и можно было бы как-то исправить, но совершенно не видно как.

Вы видите, что и из физических соображений нужно требовать, чтобы совокупность преобразований теории относительности составляла группу. Это все-таки удивительно (или не удивительно, смотря по тому, как на это смотреть): получая лоренцовы преобразования, мы вовсе не заботились о том, будут ли они группой. Мы

получили их из определенных требований: физически — из двух постулатов, математически — из условия, чтобы они оставляли инвариантным определенное выражение,— и как добавку получили, что если это так, то они образуют группу. Конечно, связь между тем, что совокупность образует группу, и тем, что определенное выражение является инвариантным, как вы уже видите, не случайна,— это глубокая связь. Но, как бы то ни было, можно поставить и обратный вопрос, и вопрос действительно был так поставлен.

Прежде чем к этому обратиться, я хочу вам посоветовать в качестве простого примера и хорошего упражнения сделать следующее. Найдите произведение двух специальных лоренцовых преобразований и покажите, что это произведение действительно принадлежит совокупности, т. е. получите это не из понятия инвариантности, а из того, что

$$A = \begin{pmatrix} \frac{1}{\sqrt{1-\beta_1^2}} & -\frac{\beta_1}{\sqrt{1-\beta_1^2}} \\ -\frac{\beta_1}{\sqrt{1-\beta_1^2}} & \frac{1}{\sqrt{1-\beta_1^2}} \end{pmatrix}, \quad B = \begin{pmatrix} \frac{1}{\sqrt{1-\beta_2^2}} & -\frac{\beta_2}{\sqrt{1-\beta_2^2}} \\ -\frac{\beta_2}{\sqrt{1-\beta_2^2}} & \frac{1}{\sqrt{1-\beta_2^2}} \end{pmatrix}.$$

Найдите преобразование, которое получается в результате последовательного произведения обоих. Для этого нужно перемножить матрицы. Произведя последовательно два преобразования, вы получите матрицу

$$C = BA = \frac{1+\beta_1\beta_2}{\sqrt{1-\beta_1^2}\,\sqrt{1-\beta_2^2}} \begin{pmatrix} 1 & -\frac{\beta_1+\beta_2}{1+\beta_1\beta_2} \\ -\frac{\beta_1+\beta_2}{1+\beta_1\beta_2} & 1 \end{pmatrix}.$$

В таком виде это выражение несколько непонятно: здесь стоит матрица и перед ней число. Условимся считать, что если перед матрицей стоит число, то каждый член матрицы умножается на это число. Совсем не видно сразу, действительно ли это такого же вида матрица, только с каким-то новым значением параметра β_3. Попробуйте доказать, что это действительно такая же матрица, и найдите, как выражается β_3 в функции от β_1 и β_2. Может быть, вы и сразу угадаете, как будет выражаться β_3. А теперь вернемся к поставленному вопросу.

Фактически было так, что физика дала два постулата, в результате которых пришли к лоренцовым преобразованиям, т. е. эти преобразования были выведены на основе требования постоянства скорости света (независимости скорости света от движения источника) и принципа относительности. Теперь вопрос ставится иначе.

Мы видели, что лоренцовы преобразования образуют группу и что не всякие преобразования обладают этим свойством. Допустим, что мы не знаем ни о принципе относительности, ни о постоянстве скорости света. Пусть имеются линейные преобразования с одним параметром

$$\begin{aligned} x' &= a(v)\,x + b(v)\,t, \\ t' &= c(v)\,x + d(v)\,t, \end{aligned} \tag{27}$$

причем мы знаем, что при произвольных $a(v)$, $b(v)$, $c(v)$ и $d(v)$ совокупность группой не будет. Мы видели, однако, что свойство быть группой жизненно для физики. Что же можно сказать об этих преобразованиях, если с самого начала потребовать, чтобы они образовывали группу? Интересно выяснить, какие ограничения налагает это требование, насколько оно суживает общий вид таких преобразований, какие формулы мы должны получить уже вследствие одного этого условия. Это чисто математическая задача: имеется совокупность линейных однородных преобразований с одним параметром. Спрашивается, когда такая совокупность является одночленной группой? В таком виде вопрос был поставлен и разрешен Франком и Роте.

Мне хотелось бы рассказать вам здесь об основной теореме теории непрерывных групп — теореме Софуса Ли, которому мы обязаны разработкой теории непрерывных групп и применением их к геометрии. Вы знаете, что одно из сильных течений, имеющихся в геометрии, опирается на утверждение, что вся геометрия сводится к теории групп. Здесь существует много остроумных положений, но начал все это как раз не математик, а Гельмгольц в своей работе «Was ist der Ursprung der geometrischen Axiome». У него это было сделано математически не совсем чисто. Ли это выяснил, и у Ли все получается чрезвычайно красиво. Одна теорема, о которой я буду говорить, позволит нам решить и поставленный нами вопрос.

У нас есть линейные уравнения (27), в которых v — произвольно меняющийся параметр. Сказать, какими функциями должны быть a, b, c, d, чтобы это было группой, не так просто, но вот что говорит Ли. Составим дифференциальные уравнения этих преобразований; оказывается, что относительно этих дифференциальных уравнений очень просто сказать, какими свойствами они должны обладать для того, чтобы это было группой. Что значит составить дифференциальные уравнения? Будем рассматривать x' и t' как функции от v при постоянных x и t

$$\begin{aligned} x' &= f_1(x, t, v), \\ t' &= f_2(x, t, v). \end{aligned} \tag{28}$$

Согласно Ли, можно рассматривать дело так, что уравнения (27) есть интегралы дифференциальных уравнений для x' и t' по v, а x и t — произвольные постоянные. Мы можем продифференцировать эти два уравнения (28) по v при постоянных x и t. У нас получится четыре уравнения, из которых мы исключим x и t (так ведь и учат составлять дифференциальные уравнения, когда впервые начинают о них рассказывать) и получим уравнения

$$\frac{dx'}{dv} = \varphi_1(x', t', v), \qquad \frac{dt'}{dv} = \varphi_2(x', t', v). \tag{29}$$

Интегрируя эти уравнения, мы предположим, что для $v = 0$ должно получаться тождественное преобразование, т. е. начальные условия должны быть такими:

$$\left.\begin{aligned} x' &= x, \\ t' &= t, \end{aligned}\right\} \quad \text{при } v = 0. \tag{30}$$

Итак, мы можем теперь рассматривать преобразования (28) как интеграл дифференциальных уравнений (29) при начальных условиях (30). Ли показал, что если преобразования (28) представляют собой группу, то дифференциальные уравнения группы (29) должны быть совершенно определенными, должны иметь вполне типичную форму. В этом все дело. Значит, нужно показать, какие ограничения налагаются на эти дифференциальные уравнения для того, чтобы их интегралы были группой.

В следующий раз мы это рассмотрим и покажем, как красиво вытекают свойства преобразований из одного лишь требования групповости.

ЧЕТЫРНАДЦАТАЯ ЛЕКЦИЯ

(14.VI 1934 г.)

Краткое резюме. Основная теорема теории непрерывных групп Ли. Существование аддитивного параметра. Ограничение вида искомого преобразования требованиями групповости и выполнения принципа относительности. Два типа преобразований — лоренцово и поворот осей. Заключение

В прошлый раз мы занимались изучением свойств, принадлежащих лоренцовым преобразованиям, как совокупности. Мы видели, что эта совокупность образует группу и что это естественно получается из требований линейности преобразований и существования у них инварианта.

Мы видели далее, что свойство быть группой физически чрезвычайно существенно, что без этого все построение физики было бы чрезвычайно затруднительным и едва ли могло бы привести к чему-нибудь хорошему. Общее лоренцово преобразование относится к группе, содержащей 10 параметров. Такую группу обозначают обычно так: G_{10}. Однородное лоренцово преобразование является подгруппой этой группы и имеет 6 параметров, т. е. это G_6; наконец, специальное лоренцово преобразование характеризуется одним параметром, т. е. это G_1.

Я предложил вам в качестве примера проверить, что специальное лоренцово преобразование действительно представляет группу. В самом деле, возьмите лоренцово преобразование со скоростью v_1. Соответствующая матрица будет L_{v_1}. Она дает нам переход от нештрихованных координат к штрихованным; затем матрица L_{v_2} дает переход от координат с одним штрихом к координатам с двумя штрихами. Если мы последовательно применим оба преобразования, то получим преобразование нештрихованных координат к координатам с двумя штрихами, имеющее матрицу

$$L_{v_3} = L_{v_2} \cdot L_{v_1}.$$

Нетрудно убедиться, что матрица L_{v_3} построена совершенно так же, как L_{v_1} и L_{v_2}, причем значение параметра в ней есть

$$v_3 = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}.$$

Это не что иное, как эйнштейновская формула сложения скоростей. Так оно и должно было получиться. Мы имели систему K', движущуюся со скоростью v_1 по отношению к системе K, и другую систему K'', движущуюся со скоростью v_2 по отношению к системе K'. Какова скорость системы K'' по отношению к системе K'? По Эйнштейну, если мы сложим эти скорости, то результирующее преобразование должно иметь в качестве параметра написанное значение.

Вопрос, который мы затем поставили, можно формулировать следующим образом. Для того чтобы вывести лоренцово преобразование, мы опирались на несколько предпосылок. Во-первых, мы требовали, чтобы определенная квадратичная форма при переходе от одной системы к другой оставалась инвариантной. Это было главное условие, причем пока мы требовали только постоянства скорости света, то оно давало равенство формы нулю. Когда же мы присоединили и требование относительности, то получили инвариантность. Теперь мы хотим поставить несколько формальный вопрос — обратный. Мы знаем, что переход от одной движущейся системы

к другой происходит при помощи линейного преобразования, причем параметром служит скорость. Мы имеем, таким образом, совокупность линейных преобразований и хотим узнать, каковы эти преобразования. Чтобы их определить, нам нужно поставить какие-то требования. Мы не будем говорить о постоянстве скорости света, не будем говорить об инвариантности какой-то формы, но потребуем, чтобы эти преобразования представляли собой группу. Какие ограничения налагает это требование на преобразования? Мы видели, что само по себе оно очень существенно, но еще не определяет вида преобразования однозначно. Посмотрим, какие же требования нужно добавить, чтобы получить вполне определенные — лоренцовы — преобразования. Для решения этого вопроса мы прежде всего должны знать, какие требования нужно предъявить аналитически, если мы хотим, чтобы преобразование было группой.

Пусть наше преобразование задается пока не в виде линейных, а в виде любых функций с одним параметром

$$x' = f_1(x, y; a), \quad y' = f_2(x, y; a). \tag{31}$$

По определению группы, нужно требовать следующего. Перейдем от x', y' к x'', y'' с помощью преобразования

$$x'' = f_1(x', y'; b),$$
$$y'' = f_2(x', y'; b),$$

принадлежащего к той же совокупности. Мы можем затем подставить x', y' в это второе преобразование и получим x'', y'' как функции от x, y, a, b. Это будут, вообще говоря, какие-то новые функции. Но если окажется, что эти функции можно опять выразить так:

$$x'' = f_1(x, y; c),$$

$$y'' = f_2(x, y; c),$$

причем

$$c = \psi(a, b),$$

то, значит, новое преобразование относится к той же совокупности и последняя образует группу.

Я уже говорил в прошлый раз, что указать сразу, какие свойства должны иметь функции f_1 и f_2, трудно. Ли пошел по другому пути. Он говорит: посмотрим на эти функции как на интегралы некоторых дифференциальных уравнений по отношению к a, для которых x, y будут начальными значениями. (Мы предполагаем, что функции дифференцируемы, непрерывны и вообще с ними можно де-

лать то, что нам нужно.) Напишем эти дифференциальные уравнения

$$\frac{dx'}{da} = \psi_1(x', y', a),$$
$$\frac{dy'}{da} = \psi_2(x', y', a), \tag{32}$$

причем при $a = a_0$, $x' = x$, $y' = y$. И вот оказывается, что здесь имеет место одна теорема, являющаяся основной теоремой теории непрерывных групп Ли (я ее формулирую в упрощенном виде — только для двух переменных и для одночленных групп). Эта замечательная теорема очень просто доказывается[1], и гласит следующее: если преобразования (31) образуют группу, то дифференциальные уравнения (32) имеют вполне определенный вид, а именно:

$$\left.\begin{aligned} \frac{dx'}{da} &= \lambda(a)\, X(x', y'), \\ \frac{dy'}{da} &= \lambda(a)\, Y(x', y'). \end{aligned}\right\} \tag{33}$$

Это необходимое и достаточное условие, причем функция $\lambda(a)$ в обоих уравнениях одна и та же. Она определена не совсем однозначно в том смысле, что ее можно умножить на произвольное постоянное число, т. е. функция от a определена с точностью до произвольного постоянного множителя. Это несущественно, в общем она все-таки однозначна и для данной группы является вполне определенной функцией a.

Мы примем эту теорему без доказательства и посмотрим, какие следствия можно из нее вывести. А следствия можно вывести довольно интересные. Предположим, что написанные выше преобразования (31) образуют группу. Мы можем заменить a каким-нибудь другим параметром, например $\alpha = \alpha(a)$. Это не изменит группы. Если это группа при параметре a, то при a это также будет группой, т. е. в качестве нового параметра можно взять любую функцию старого. В лоренцовых преобразованиях мы считали параметром скорости, но могли бы считать параметром и скорость в квадрате или скорость в квадрате плюс единица. В этом легко убедиться.

Далее мы видели, что произведение двух преобразований группы дает новое преобразование, соответствующее новому значению параметра, которое, вообще говоря, не является суммой для первых значений (например, в лоренцовых преобразованиях это не сумма).

---

[1] См., например, *Гурса*. Курс анализа, т. II, ч. II, стр. 94.

Введем теперь новый, более удобный параметр

$$\varphi = \int_{a_0}^{a} \lambda(a)\, da \tag{34}$$

(при $a = a_0$ он равен нулю).

Если мы подставим этот параметр в дифференциальное уравнение (33), то получим

$$\begin{aligned} \frac{dx'}{d\varphi} &= X(x',\ y'),\\ \frac{dy'}{d\varphi} &= Y(x',\ y'), \end{aligned} \tag{35}$$

т. е. при этом параметре уравнения группы должны приобрести такой вид, что в правых частях параметра нет. Это типично для дифференциальных уравнений группы, и это имеет место только при специальном выборе параметра (34). Всегда существует один такой параметр, который приводит дифференциальные уравнения к виду (35). Отсюда получается очень интересное следствие. Дифференциальные уравнения (35) можно интегрировать в том смысле, что можно указать свойства общего интеграла. В более привычной форме эти уравнения записываются так:

$$\frac{dx'}{X} = \frac{dy'}{Y} = d\varphi. \tag{35'}$$

Вы знаете, что первый интеграл таких уравнений есть

$$\Omega_1(x',\ y') = C_1. \tag{36}$$

Я могу взять далее одно из уравнений (35′), подставить (36), проинтегрировать и, заменяя C_1 опять при помощи выражения (36), получу второй интеграл

$$\Omega_2(x',\ y') = C_2 + \varphi. \tag{37}$$

Интегрирование всегда приводит к таким выражениям [1].

Я могу теперь выразить C_1 и C_2 при помощи начальных ($\varphi = 0$) значений x и y. В результате я получу следующие интегралы уравнений группы:

$$\left.\begin{aligned} \Omega_1(x',\ y') &= \Omega_1(x,\ y),\\ \Omega_2(x',\ y') &= \Omega_2(x,\ y) + \varphi. \end{aligned}\right\} \tag{38}$$

Значит, если есть группа, то при определенном выборе параметра ее дифференциальные уравнения непременно имеют вид (35′).

---

[1] См. *Гурса*, loc. cit.

Эта форма уравнений называется канонической. Интегралы этих уравнений имеют вид (38).

Обратно, если даны уравнения вида (35′), то интегралы всегда будут группой, т. е. уравнения (35′) всегда представляют группу. Сейчас мы это докажем.

Я хочу перейти от x' к x'' и пишу

$$\left.\begin{aligned}\Omega_1(x'', y'') &= \Omega_1(x', y'),\\ \Omega_2(x'', y'') &= \Omega_2(x', y') + \varphi_2.\end{aligned}\right\} \tag{39}$$

Произведение этих двух преобразований дает переход от x, y к x'', y''. Подставив (38) в (39), получаем

$$\left.\begin{aligned}\Omega_1(x'', y'') &= \Omega_1(x, y),\\ \Omega_2(x'', y'') &= \Omega_2(x, y) + \varphi_1 + \varphi_2.\end{aligned}\right\}$$

Таким образом, мы перешли от нештрихованной к дважды штрихованной системе координат и получили такую же совокупность уравнений, но с другим значением параметра. В первом случае у нас был параметр φ_1, во втором φ_2; произведение этих двух преобразований имеет тот же вид, но с параметром

$$\varphi_3 = \varphi_1 + \varphi_2. \tag{40}$$

Этот результат — аддитивность — мы получили сверх всякого ожидания. Мы исходили из того факта, что если наши преобразования образуют группу, то дифференциальные уравнения имеют вполне определенный вид. Из этого вида мы пришли к заключению, что всегда можно ввести такой параметр, при котором последовательное применение двух преобразований соответствует третьему преобразованию, параметр которого есть просто сумма параметров первого и второго преобразований. Оказывается, что так подобрать параметр всегда можно, причем он определяется однозначно, с точностью до численного коэффициента (т. е. если вместо φ мы возьмем $k\varphi$, то получим то же самое).

Всякая одночленная группа может быть при помощи соответствующего подбора параметра приведена к такой форме, в которой параметры просто складываются.

Между прочим, это сразу наводит на следующий вопрос. Мы видели, что когда за параметр лоренцова преобразования мы принимаем скорость, то произведение двух преобразований имеет параметр, который не равен сумме параметров. Мы можем сказать теперь, что есть какой-то один параметр, при котором в произведении двух лоренцовых преобразований мы, наверно, получим простую сумму параметров. Скорость не есть такой параметр, но такой параметр непременно существует. К этому мы обратимся позднее.

Сейчас нас, в сущности, интересует такой вопрос: что можно сказать относительно группы при требовании, чтобы преобразования были линейными и однородными?

Будем вести наши исследования, принимая такой параметр, который просто складывается,— так будет гораздо проще. Потом мы свяжем этот параметр со скоростью и увидим, какие следствия отсюда получаются. Основная теорема была приведена без доказательства. Мы вывели из нее определенные следствия и сейчас применим их к интересующему нас случаю и посмотрим, что из этого следует.

Самый общий вид линейного преобразования с одним параметром (будем пока считать параметром скорость) такой:

$$\left.\begin{aligned} x' &= a(v)\,x + b(v)\,t, \\ t' &= c(v)\,x + d(v)\,t, \end{aligned}\right\} \tag{41}$$

где a, b, c, d — произвольные функции. Одно только можно сказать сразу же: если мы хотим, чтобы параметр обозначал скорость, то должна иметь место некоторая связь между коэффициентами, а именно $\frac{b(v)}{a(v)} = -v$, которая следует просто из физического смысла скорости.

Еще одно замечание. Из лоренцова преобразования мы заключили, что длина, измеренная в одной системе, иная, чем измеренная в другой. Из всего хода доказательства видно, что отношение длин данного стержня в одной и в другой системе определяется коэффициентом $a(v)$. Таков физический смысл этого коэффициента; заметим это, хотя пока это и неважно.

Мы хотим найти, какую форму должны иметь преобразования (41), чтобы они были группой. Воспользуемся нашим рецептом: продифференцируем уравнения (41) по v и получим дифференциальные уравнения

$$\frac{dx'}{dv} = \alpha(v)\,x' + \beta(v)\,t',$$

$$\frac{dt'}{dv} = \gamma(v)\,x' + \delta(v)\,t'.$$

Если это группа, то дифференциальные уравнения должны иметь вид

$$\frac{dx'}{dv} = \lambda(v)\,(\alpha x' + \beta t),$$

$$\frac{dt'}{dv} = \lambda(v)\,(\gamma x' + \delta t'),$$

т. е. справа могут стоять только такие однородные линейные функ-

ции x' и t', в которых α, β, γ, δ — постоянные величины. Введем новый параметр

$$\varphi = \int_0^v \lambda(v)\, dv. \tag{42}$$

Тогда канонические уравнения искомых преобразований должны иметь вид

$$\begin{aligned} \frac{dx'}{d\varphi} &= \alpha x' + \beta t', \\ \frac{dt'}{d\varphi} &= \gamma x' + \delta t'. \end{aligned} \tag{43}$$

То, что α, β, γ, δ — постоянные коэффициенты,— это колоссальное ограничение, которое получилось именно из требований, чтобы уравнения представляли группу. Проинтегрировать уравнения (43) ничего не стоит, и мы могли бы прямо в окончательной форме написать линейное однородное преобразование, удовлетворяющее только тому условию, что оно образует группу. И все-таки это будет набор совокупностей, каждой из которых будут отвечать определенные значения α, β, γ, δ. Какие это числа, мы не знаем. Условие группы требует только, чтобы это были постоянные величины. Можно взять числа 1, 2, 3, 4 или 1, —2, —3, —4. Типы преобразований будут существенно различны в зависимости от соотношений между этими коэффициентами и это приводит нас к изучению возможных типов преобразований. Уравнения (43) хорошо известны и подробно продискутированы, и на этом останавливаться мы не будем. Хотя и интересно с точки зрения только условия группы рассмотреть, как выглядят отдельные преобразования в зависимости от различных соотношений между числами α, β, γ, δ, и посмотреть, каким физическим смыслом каждое из них может обладать, но мы пойдем не таким путем. Мы сразу же поставим дальнейшие требования, потому что мы видели, что при выводе лоренцова преобразования мы также не могли удовлетвориться одним только постулатом постоянства скорости света, а требовали еще и относительности. Поэтому мы постараемся и здесь, прежде чем интегрировать уравнения, ввести это требование.

Как выражается требование относительности? Возвратимся к исходным уравнениям (41). Мы можем их обратить и получим тогда

$$x = \frac{dx' + bt'}{ad - bc}, \qquad t = \frac{-cx' + at'}{ad - bc}. \tag{44}$$

Если система K' движется по отношению к системе K со скоростью v, то я должен требовать, чтобы система K двигалась по отношению к

системе K' со скоростью $-v$. Это первое требование. Второе требование, которое мы предъявляли и раньше, таково: если я измерил неподвижный в K' масштаб в системе K и он оказался в n раз короче, то и обратно, когда я измеряю из системы K масштаб, который неподвижен в системе K, я должен получить такое же укорочение.

Эти требования налагают на наши коэффициенты довольно большие новые ограничения, а именно: во-первых, скорость, обратная скорости $v = -\frac{b(v)}{a(v)}$, согласно (44), есть $v' = \frac{b(v)}{d(v)}$. Я требую, чтобы $v' = -v$, откуда вытекает, что $a(v) = d(v)$. Кроме того, взаимное укорочение должно быть одинаковым. В (41) оно дается коэффициентом $a(v)$, а в (44) — коэффициентом $\frac{d(v)}{ad - bc}$. Следовательно,

$$a = \frac{d}{ad - bc},$$

а так как $a = d$, то детерминант должен равняться единице. Значит, введение требования относительности дает следующее:

$$\left.\begin{aligned} a(v) &= d(v), \\ ad - bc &= 1. \end{aligned}\right\} \tag{45}$$

Это довольно большое ограничение, но соответствующие требования мы ставили и раньше. Я бы сказал грубо, что мы заменили теперь условие постоянства скорости света требованием групповости. Благодаря тому, что свойство групповости соблюдено, мы пришли к тому, что уравнения (43) должны иметь постоянные коэффициенты. Для того же, чтобы и принцип относительности был соблюден, необходимо, чтобы коэффициенты интегральных соотношений удовлетворяли условиям (45).

Легко установить, как отзываются требования (45) на коэффициентах α, β, γ, δ. Это можно сделать посредством приема, которым теория групп очень часто пользуется. Я выведу, во всяком случае, необходимое условие, которому должны удовлетворять α, β, γ, δ и которое имеет очень простой вид.

При $\varphi = 0$ $x' = x$ и $t' = t$. Перейдем от $\varphi = 0$ к очень малому $\Delta\varphi$. Мы можем написать тогда

$$x' = x + \left(\frac{dx'}{d\varphi}\right)_{\varphi=0} \Delta\varphi,$$

$$t' = t + \left(\frac{dt'}{d\varphi}\right)_{\varphi=0} \Delta\varphi.$$

Это будет, как говорят, бесконечно малое преобразование, т. е. соответствующее бесконечно малым скоростям. Но из (43) мы зна-

ем $\frac{dx'}{d\varphi}$ и $\frac{dt'}{d\varphi}$ при $\varphi = 0$, так как при этом $x' = x$, $t' = t$. Подставив, получим

$$x' = (1 + \alpha\Delta\varphi)\,x + \beta\Delta\varphi t,$$
$$t' = \gamma\Delta\varphi x + (1 + \delta\Delta\varphi)\,t.$$

Мы имеем здесь просто интегрирование в ближайшей окрестности. При любом φ нужно было бы (43) проинтегрировать, но для ближайшей окрестности можно получить преобразование так, как мы это сделали. Требования (45) справедливы при любом φ, а значит, и для $\Delta\varphi$, но, конечно, все здесь должно делаться лишь до первого порядка относительно $\Delta\varphi$. Легко видеть, что если мы хотим удовлетворить двум условиям (45), то необходимо $\alpha = \delta = 0$. Таким образом, дифференциальные уравнения наших преобразований, если прибавить требование относительности, должны быть таковы:

$$\frac{dx'}{d\varphi} = \beta t', \qquad \frac{dt'}{d\varphi} = \gamma x', \tag{46}$$

где β и γ — произвольные постоянные числа. Это уже почти единственное преобразование. В том, что оно все же не единственное,— весь интерес. Теперь остается проинтегрировать эти уравнения и получить уравнения в конечном виде, которые удовлетворяют всем нашим условиям. Но позвольте сделать некоторые замечания.

Все наши сегодняшние рассуждения формальны. То, что мы требовали групповости,— это формальное требование; то, что мы требовали относительности, также сводилось на формальное требование, чтобы некоторые коэффициенты были связаны определенным образом, а отсюда какие-то другие коэффициенты должны были равняться нулю. Физический смысл при этом не учитывался, и все это были формальные математические условия.

Рассмотрим совершенно другую задачу. Возьмем на плоскости координатную систему x, y и перейдем к новой системе x', y'. Пусть угол между осями равен φ. Тогда мы знаем, что

$$x' = x\cos\varphi + y\sin\varphi, \tag{47}$$
$$y' = -x\sin\varphi + y\cos\varphi.$$

По отношению к этой задаче я мог бы поставить совершенно такие же вопросы. Я знаю, что уравнения преобразования линейны и однородны, и я потребую, чтобы они были группой (между прочим, на этом строится вся геометрия). Я нарочно выписал эти преобразования. Если вы посмотрите на них, то убедитесь, что два других требования здесь также выполнены. Во-первых, $a = d$ и, во-вторых, детерминант равен единице, так что постулат относи-

тельности удовлетворен. Я мог бы, повторяю, решить задачу о преобразовании координат на плоскости, исходя из следующих требований. Преобразования должны быть линейными, однородными, должны представлять группу, а кроме того, $a = d$, и детерминант должен равняться единице. Как геометрия поступает обычно? Линейное преобразование однородно, и требуется, чтобы $x^2 + y^2$ было инвариантом. Отсюда приходят к (47). Но можно поступить и обратно: ничего не говорить об инварианте, а требовать групповости. Легко показать, что это означает геометрически. Это совершенно то же самое, что наше требование взаимности сокращения масштаба в штрихованных и нештрихованных координатах.

Пока что у нас другая задача, но если я указанным образом поставлю вопрос о разыскании преобразований, то я предъявлю те же самые требования, что и в задаче о преобразовании от одной движущейся системы к другой. Совершенно естественно, что по отношению к дифференциальным уравнениям ответ должен быть тем же самым, т. е. и для поворота координат на плоскости дифференциальные уравнения преобразования должны быть такими же. Таким образом, наши требования пока что не делают различия между этими двумя физически совершенно разными задачами. Одна — преобразование координат, движущихся друг относительно друга систем отсчета, а другая — поворот осей на плоскости, т. е. преобразование двух пространственных координат. И вот интересно, что наши требования действительно содержат ответ как на первый, так и на второй вопрос, т. е. содержат оба эти случая, но никаких других случаев не содержат. Потом мы выясним, как провести различие между этими двумя случаями, но сейчас вы видите, насколько они близки. Вот почему при лоренцовых преобразованиях говорят о геометризации. Действительно, задача, поскольку она ставится абстрактно математически, та же самая, *но только математически*. Физически мы толкуем x, y и t совершенно различно: в первом случае x и y — пространственные координаты на плоскости, во втором случае x — пространственная координата, а t — время, которое совершенно иначе измеряется и определяется. Ясно, что должно быть и какое-то формальное различие, и мы увидим сейчас, что оно действительно имеется.

В чем оно заключается? Я сказал, что β и γ — произвольные постоянные. И вот оказывается существенным, имеют ли они одинаковые знаки или различные. Рассмотрим первый случай — β и γ имеют одинаковые знаки — и в этом предположении проинтегрируем наши совсем простые уравнения. Заметим, что $\sqrt{\beta\gamma}$ действителен, так как под корнем стоит положительная величина. β и γ определены с точностью до постоянного коэффициента,

и мы можем поэтому считать, что $\sqrt{\beta\gamma} = 1$. Обозначим далее $\sqrt{\gamma/\beta} = \rho$. Мы получаем тогда из (46) следующие окончательные выражения:

$$\left.\begin{aligned} x' &= \frac{e^{\varphi}+e^{-\varphi}}{2}\,x - \rho\,\frac{e^{\varphi}-e^{-\varphi}}{2}\,t,\\ t' &= -\frac{e^{\varphi}-e^{-\varphi}}{2\rho}\,x + \frac{e^{\varphi}+e^{-\varphi}}{2}\,t. \end{aligned}\right\} \tag{48}$$

Мы убедимся, что это лоренцово преобразование.

Параллельно рассмотрим второй случай, когда $\beta\gamma < 0$, т. е. когда β и γ имеют разные знаки. Мы всегда можем умножить их на такую величину, чтобы получать $\sqrt{\beta\gamma} = i$. Далее, обозначим $\sqrt{\gamma/\beta} = -i\rho$. Если проинтегрировать уравнения (46) при этих условиях, то мы получим

$$\left.\begin{aligned} x' &= \frac{e^{i\varphi}+e^{-i\varphi}}{2}\,x - \rho\,\frac{e^{i\varphi}-e^{-i\varphi}}{2i}\,t,\\ t' &= \frac{e^{i\varphi}-e^{-i\varphi}}{2\rho i}\,x + \frac{e^{i\varphi}+e^{-i\varphi}}{2}\,t. \end{aligned}\right\} \tag{49}$$

Я пока не знаю, какие уравнения отвечают моему случаю. Я утверждаю, что первые есть лоренцово преобразование, а вторые — самое обычное преобразование координатной системы. Но пока я не знаю ни лоренцовых преобразований, ни этих, и я спрашиваю: можно ли именно из физических соображений, из того, что определение времени иное, чем определение второй координаты в пространстве, сказать с самого начала, какие уравнения отвечают одному случаю и какие другому? Я утверждаю, что можно, и притом различными способами. Я это сделаю только одним способом.

Возьмем уравнения (48). Форма этих уравнений, которую мы получили из требования групповости, сразу говорит о том, что существует инвариант. Возведите первое и второе уравнения в квадрат, умножьте второе на ρ^2 и вычтите из первого. Вы получите тогда

$$x'^2 - \rho^2 t'^2 = x^2 - \rho^2 t^2,$$

т. е. это выражение является инвариантом. Аналогично во втором случае инвариантом является

$$x'^2 + \rho^2 t'^2 = x^2 + \rho^2 t^2,$$

причем оба раза ρ — действительная величина. Итак, уравнения, которые получились из наших требований, говорят, что в обоих

случаях существует инвариант, но в первом случае — один, во втором — другой.

Отсюда я сразу могу сказать, какой случай относится к движению, а какой к преобразованию координат. По определению времени и движения, я могу сказать следующее: если у вас есть неподвижная точка в какой-то координатной системе и если эта система движется, то с ростом времени координата, отсчитанная в неподвижной системе, несомненно будет возрастать. Это содержится в самих наших понятиях о времени и о движении. Значит, если существует инвариант между x и t, то он может быть только таким, что при росте t должно расти и x. Если же t и x растут и при этом какое-то выражение остается инвариантным, то это может быть только выражение со знаком минус. В пространстве наоборот — поворот осей не изменяет расстояния между двумя точками. Если x растет до бесконечности или t растет, то вся величина растет. Но при изменении координатной системы требования, чтобы x и t увеличивались, нет.

Таким образом, из тех требований, которые мы предъявляем ко времени и движению, ясно, что для нашего случая инвариантным может быть только первое выражение, т. е. мы однозначно выделили теперь преобразование, которое нас интересует.

В этом «небольшом» различии в знаке кроется сущность различия между пространственной координатой во втором случае и временем, как другой «координатой», в первом случае.

Нам осталось показать, что (49) действительно есть не что иное, как обычное преобразование. Легко видеть, что уравнения (49) можно переписать в виде

$$x' = x \cos \varphi - \rho t \sin \varphi,$$

$$t' = -x \frac{\sin \varphi}{\rho} + t \cos \varphi,$$

а это обычная форма преобразования, за исключением произвольной величины ρ. Но ее присутствие понятно. Мы ведь не требовали, чтобы масштаб для x и y был одинаков. При этом добавочном требовании мы получим обычные формулы преобразования координат.

Наоборот, такая форма совершенно непривычна для нас в лоренцовом преобразовании. Случайна ли эта форма или не случайна? Я уже указал, что параметр φ — это специально подобранный параметр, который обладает тем свойством, что умножению двух преобразований соответствует сложение их параметров. Так уж повелось, что при преобразовании осей за параметр принимается угол φ, и когда вы поворачиваете оси сначала на угол φ_1,

а затем на угол φ_2, то окончательный поворот будет на угол $\varphi_3 = \varphi_1 + \varphi_2$. Но можно было бы выбрать в качестве параметра для поворота осей не φ, а, скажем, $\operatorname{tg}\varphi$. Пусть $\operatorname{tg}\varphi = v$. Если подставить этот параметр, то получится лоренцово преобразование. Действительно, $\varphi_3 = \varphi_1 + \varphi_2$, а значит, $v_3 = \operatorname{tg}\varphi_3 = \operatorname{tg}(\varphi_1 + \varphi_2)$, т. е.

$$v_3 = \frac{\operatorname{tg}\varphi_1 + \operatorname{tg}\varphi_2}{1 - \operatorname{tg}\varphi_1 \operatorname{tg}\varphi_2} = \frac{v_1 + v_2}{1 - v_1 v_2}.$$

Но это как раз лоренцово преобразование для скоростей. Таким образом, если бы поворот характеризовался не самим углом, а его тангенсом, то мы получили бы лоренцовы преобразования. Мы можем и движение одной системы по отношению к другой характеризовать не скоростью v, а величиной φ. Тогда мы получим аддитивность сложения этих новых «скоростей». Такое φ в качестве физического параметра вместо скорости именно с этой точки зрения было введено Робом. Роб говорит следующее: мы можем охарактеризовать движение одного тела по отношению к другому скоростью. Что такое скорость? Мы знаем, что тело может двигаться скорее или медленнее. В качестве меры мы избрали отношение пути ко времени и приняли эту величину в качестве параметра. Ничто не мешает нам выбрать другую характеристику движения. Позвольте не останавливаться на этом подробно, но можно было бы физически сделать так. Представьте себе, что в некоторой системе отсчета мы стреляем из пушки; ядро будет двигаться, и это движение мы примем за меру скорости. Скорость какого-то другого заданного движения мы можем определить таким образом: мы выстрелили из первой пушки; вторую, совершенно такую же пушку мы заставляем двигаться так, чтобы она шла вместе с ядром первой, и в это же время стреляем из второй пушки. Тогда ее ядро будет как-то двигаться. И вот, по определению, мы скажем, что это второе ядро имеет быстроту, вдвое большую, чем первое. Если мы так определим, то это и будет наш аддитивный параметр. Мы можем назвать его *быстротой.* (Роб назвал его rapidity). Это был бы очень ценный параметр именно в силу своей аддитивности.

Теперь позвольте показать, что (48) действительно лоренцово преобразование. В нем стоят не круговые, а гиперболические функции φ, т. е. его можно записать так:

$$x' = \operatorname{ch}\varphi\; x - \rho \operatorname{sh}\varphi\; t,$$

$$t' = -\frac{\operatorname{sh}\varphi}{\rho}\, x + \operatorname{ch}\varphi(t).$$

Если мы обозначим $\rho \operatorname{th}\varphi = v$ (а мы обязаны так сделать, если хотим обычным образом определить скорость v как путь, деленный на

время), то уравнения примут такой вид:

$$\left.\begin{aligned} x' &= \frac{x - vt}{\sqrt{1 - \frac{v^2}{\rho^2}}}, \\ t' &= \frac{t - \frac{v}{\rho^2}x}{\sqrt{1 - \frac{v^2}{\rho^2}}}, \end{aligned}\right\} \qquad (50)$$

а это и есть наше обычное лоренцово преобразование. Единственное, чего наши требования не определяют,— это величины ρ. Для этого нужны физические соображения, нужен какой-то физический процесс. Принцип относительности утверждает, что ρ — скорость света в вакууме.

Итак, мы видим, что могут дать наши требования и чего не могут. С одной стороны, они указали нам тип преобразований, но еще оставили большой произвол. Затем мы потребовали, чтобы одна из координат сообразовалась с нашим понятием времени, и мы нашли тогда лоренцово преобразование, но не нашли значения ρ. Это и понятно: из соображений группы численная величина получиться не могла. Для этого нужен опыт.

Можно было бы поступить несколько иначе: сразу ввести требование относительности, т. е. наложить условия (45) и затем ввести требование групповости. Это было сделано Франком. Предположение, что скорость одной системы по отношению к другой обратна скорости второй по отношению к первой и что имеет место одинаковое взаимное сокращение масштаба, достаточно, чтобы привести преобразование к форме (50). Но тогда зачем нужно требование групповости? Я сказал, что преобразования приводятся к форме (50), но с одной разницей: ρ остается произвольной функцией от скорости v. Требование групповости приводит к тому, что ρ — постоянная величина, а это является решающим.

Вы видите, таким образом, как различные требования, одни из которых кажутся большими, другие маленькими, налагают известные ограничения и как отсюда получаются определенные соотношения. Кроме того, указать на групповой характер преобразований важно, как мне кажется, и потому, что становится ясным значение φ. В книгах часто пишут так, что просто сближают преобразование координат на плоскости и лоренцово преобразование. Мне кажется, что если рассматривать с указанной точки зрения, то все получает несколько другой аспект. Замечу еще, что если вместо x написать x_1, а вместо ict взять x_4 (причем то, что я ввожу, не имеет никакого ни мистического, ни физического значения, а просто вводится мнимая координата), то для этих коор-

динат лоренцово преобразование приобретает совершенно такой же вид, как для обычного поворота осей, но только φ — мнимый угол. Опять-таки это ничего не значит физически, это только способ обозначения. Часто говорят поэтому, что лоренцово преобразование есть поворот осей, но на мнимый угол.

Гораздо сложнее применить групповые соображение к общим лоренцовым преобразованиям G_6 и G_{10}. Я не буду на этом останавливаться. Укажу только, что и в обычной геометрии все, что можно отсюда вывести, представляет собой чрезвычайно занимательные и фундаментальные вещи. Теперь этот круг вопросов очень хорошо разработан, а начало его было положено знаменитым мемуаром «О происхождении геометрических аксиом» Гельмгольца, который ко всем наукам подходил как естествоиспытатель. Хотя Гельмгольц и сделал довольно грубую логическую ошибку, исправленную потом Ли, но, несмотря на это, творцом всей постановки вопроса нужно считать именно его.

Можно было бы еще много говорить об этом, но нужно закончить. Позвольте в заключение сказать еще несколько слов.

Я особенно старался подчеркнуть (я думаю, что это мое желание и самое мое мнение вы ясно усвоили), что все вопросы, связанные с теорией относительности,— прежде всего вопросы физические. Ее философское значение — это особый вопрос, нас же интересовало физическое значение. Это важно подчеркнуть вот почему: физическая теория оперирует некоторыми определениями, но такими определениями, в основе которых лежат действительно выполнимые физические процессы в реальной природе. Если я говорю, что я измеряю длину, то я требую, чтобы измерение производилось не при помощи каких-то мифических тел, а с тем или иным приближением при помощи физических тел, чтобы можно было указать, *как* это сделать. Если я говорю о времени, о синхронности часов здесь и там, то я требую, чтобы указали, как это установить. *В зависимости от того, что́ показывает эксперимент, и наши требования и наши определения должны быть различны.* Если мы уверены, если все опыты говорят, что в природе не существует сигнальных скоростей, бо́льших скорости света, то, определяя, скажем, одновременность, синхронизм двух часов в различных местах, мы должны с этим считаться. Мы не можем говорить, что определим одновременность, пустив мгновенный сигнал. Мы должны считаться с тем, что есть, и если такого сигнала не существует, то нельзя его ставить в основу теории. Если это так и если признать опыты, с которыми мы соглашаемся, если признать опыт Майкельсона правильным, то те предпосылки, из которых исходит теория относительности, ни в коей мере не являются более искусственными, чем те, из которых исходила классика. Классика тоже должна была опираться на предположения и гипотезы. Классика

также должна была определять скорость и устанавливать синхронизм часов здесь и в другом месте. А тогда я могу от представителя классической физики требовать: скажи, как ты устанавливаешь синхронизм?

Если он говорит, что посылает сигнал с бесконечной скоростью, то я это отвергаю, потому что физика должна считаться с тем, что есть в природе. Он может сказать иначе: я переношу часы в другое место и считаю, что они синхронны. Тогда вопрос касается факта: как ведут себя часы, которые таким образом переносят? Релятивист говорит, что они ведут себя так-то, классик говорит, что так-то. Никакой принципиальной разницы в постановке вопроса нет. Все существующие опытные данные говорят, что правы те, кто придерживается теории относительности. Я подчеркиваю еще раз, что не априорные соображения заставляют принять эту теорию, а то, что она более естественна, чем классика, что она оперирует определенными реальными экспериментами. Если эти опыты неправильны, то и теория относительности неправильна, но исходить надо из определенных опытов. Если исходить из тех опытов, о которых мы говорили, то теория относительности не искусственна, а естественна. В ней нет какой-то надстройки над классикой, в которой будто бы все было «само собой очевидно». Если говорить о надстройках, то классика также должна была их делать, но только другие. Кто прав,— это вопрос опыта, и то, что теперь известно, говорит за теорию относительности. Один из главных аргументов — вопрос о скорости света. Признание скорости света в вакууме предельной ведет к таким последствиям, от которых физику отмахнуться нельзя.

К таким же заключениям приводят mutatis mutandis и вопросы квантовой механики. И там вопрос заключается в опыте. С теорией относительности было так, что пока имели дело с маленькими скоростями, никаких отклонений от классики не замечали. Пересмотреть всю систему заставили новые *факты*, теория относительности возникла именно отсюда, а не «откуда-то вдруг взялась». Нечто подобное (я не говорю — то же самое, потому что боюсь таких аналогий) получилось и в квантовой механике. Пока дело касалось макрокосмоса, где влияние на наблюдаемый предмет либо мало, либо может быть учтено, создавалось впечатление, что всегда и везде будет обязательно так. Но если признать опыты, и в первую очередь опыт Комптона, правильными, если признать, что имеется неустранимое взаимодействие, то подобное убеждение рушится. Квантовая механика не произвольное построение. Такого построения требует физика, и если его не сделать, то получится противоречие с опытными фактами. Конечно, всегда можно возразить две вещи: либо опыт неверен, либо можно построить еще третьим способом. Это верно, но не нужно забывать, что и

прежнее построение содержало предположения. В этом отношении разницы для физика нет — и раньше должны были делать предположения и теперь, какие предположения правильны, показывает опыт.

Нас интересовали другие вопросы, квантовая механика к ним не относится, но, заканчивая, я хотел подчеркнуть эту сторону дела. Я уверен, что каждый из вас, если не будет останавливаться на полдороге и думать, что теория относительности переворачивает его понятия, а будет трезво относиться к опыту и ко всем построениям, имея в виду, что переворачиванием понятий часто называется разрушение предрассудков, не встретит никаких трудностей, за исключением, быть может, конкретных трудных вопросов.

СЕМИНАР
ПО ИЗБРАННЫМ ВОПРОСАМ ОПТИКИ

(1934—1935 гг.)

О СВЯЗИ КЛАССИЧЕСКОЙ И КВАНТОВОЙ ТЕОРИИ ДИСПЕРСИИ

(14.VI 1935 г.)[1]

Я хочу остановиться на вещах, которые не прямо примыкают к тем вопросам, которые разбирались на семинаре, но о которых, по-моему, все же следует поговорить. Речь идет об одном вопросе, касающемся связи квантовой и классической теории прохождения света через тела. Безусловно, базой современной физики является классика, но, с другой стороны, кванты,— по-видимому, единственный правильный путь для глубокого понимания микрокосмоса. Мы возьмем прохождение света через прозрачные тела — это большой круг проблем, охватывающий дисперсию в узком смысле, керр-эффект, магнитное вращение плоскости поляризации и т. д.,— и на нем выясним, как далеко простирается сходство, в чем заключаются различия,— одним словом, какова связь классической и квантовой трактовки вопроса.

В классике исходными являются основные максвелл-лоренцовы уравнения

$$\left.\begin{aligned} \operatorname{rot}\ \mathbf{E} &= -\frac{1}{c}\,\dot{\mathbf{H}}, \\ \operatorname{rot}\ \mathbf{H} &= \frac{1}{c}\,(\dot{E} + 4\pi\rho\mathbf{v}). \end{aligned}\right\} \qquad (1)$$

Но их недостаточно для решения вопроса, нам надо знать еще связь $\mathbf{v}$ с полем. Мы ограничимся случаем газа, когда можно пренебречь взаимодействием частиц. Рассматривая среду как непрерывную, мы должны все величины в (1) считать *средними* значениями, так что

$$\rho\mathbf{v} = Ne\,\frac{d\mathbf{r}}{dt} = \frac{d\mathbf{P}}{dt},$$

---

[1] [Лекция, прочитанная на семинаре по избранным вопросам оптики. Обработка С. М. Рытова по запискам С. М. Рытова и М. А. Леонтовича.]

где **P** — поляризация. Таким образом, все сводится к отысканию **P** как функции поля или в конечном счете среднего отклонения **r**.

Мы находим связь **r** с полем из *законов классической механики*, применяемых к электрону в атоме или молекуле, причем вся классическая теория исходит из предположения *упругой* связи электрона. Таким образом, уравнение движения электрона есть

$$m \frac{d^2\mathbf{r}}{dt^2} + k\mathbf{r} = e\mathbf{E}. \tag{2}$$

Это второй этап решения задачи — механические уравнения. Из них мы получаем $\mathbf{r} = \alpha\mathbf{E}$, и дальнейший путь ясен; он дает нам для квадрата показателя преломления

$$n^2 = 1 + 4\pi N e \alpha.$$

Поскольку $\alpha = \alpha(\omega)$, постольку мы получаем объяснение дисперсии, аномальной дисперсии и т. д.

В изложенном виде теория не учитывает случая резонанса ($\omega = \omega_0$). Чтобы охватить и этот случай, надо дополнить уравнение (2) затуханием, я сказал бы *«принципиальным» затуханием*, ибо по классике колеблющийся электрон *обязательно излучает*. Это совершенно общее положение, и ни при каких условиях оно не может быть устранено. Другие причины затухания могут быть, а могут и отсутствовать, но затухание благодаря излучению есть всегда. Конечно, вдали от резонанса им можно пренебречь.

Более сложный случай — прохождение света при наличии внешнего магнитного поля $\mathfrak{H}$. Тогда вместо (2) мы будем иметь уравнение

$$m \frac{d^2\mathbf{r}}{dt^2} + k\mathbf{r} = e\mathbf{E} + \frac{e}{c} [\dot{\mathbf{r}}, \mathfrak{H}]^1. \tag{3}$$

Задача становится сложнее, но в общем все остается по-прежнему.

Как же ставится вопрос в квантовой механике? Здесь нельзя не отметить большого преимущества классики. У Лоренца подкупает исчерпывающая полнота трактовки и простота. Конечно, модель неверна и ее усложнение дела не спасает. В этом мы убедились: кванты отсюда и возникли. Но система принципов классической теории чрезвычайно последовательна и полна. Квантовая теория далеко не так стройна, и именно в дисперсии это особенно хорошо видно, поскольку в настоящее время у нас нет настоящей

[1] Конечно, действием на электрон магнитного поля **H** самой волны мы пренебрегаем: оно имеет порядок $\frac{v}{c}$ по отношению к действию электрического поля **E**.

квантовой *теории поля* (механика есть, и вполне удовлетворительная).

Современная квантовая теория дисперсии в первом приближении поступает так. Она *перенимает* уравнения поля (1) без всяких изменений из классики, а уравнение движения (2) заменяется волновой механикой. Конечно, это непоследовательно, это компромисс, но все же он дал уже много ценных результатов. На этой, пожалуй, пока что единственной стадии мы и остановимся.

Мы не будем сегодня говорить о том, почему вообще пришлось отказаться от классических уравнений движения. Самое существенное здесь не в отказе от классики, а в отказе от *модели* осциллятора. Мы знаем, что в атоме водорода хорошие результаты дает не упругая сила, а кулонов закон (это и в классике дает другие результаты). Вообще увидели, что квазиупругая модель не годна. Но, предвосхищая результат, я укажу теперь же: если работать с квазиупругим осциллятором, то в интересующих нас вопросах и классика и кванты дают *принципиально* (а не почти) *одно и то же*.

В волновой механике движение электрона определяется уравнением Шрёдингера. Этого мало — дело не в изменении уравнения, а в том, что весь подход совершенно другой. Волновая механика отказывается от разговоров о движении индивидуальной точки и имеет дело с функциями распределения, с вероятностями. В данном опыте, который я называю движением точки под действием данной силы, я имею $\Phi(x)\,dx$ — вероятность найти частицу в интервале $(x,\ x + dx)$. Если при «совершенно тех же самых» условиях повторить опыт, то классика говорит, что и результаты получатся прежние. Волновая механика утверждает, что это не так. При повторении опыта, как бы точно мы его ни повторяли, мы всегда придем только к *вероятности* $\Phi(x)\,dx$. И уравнение Шрёдингера относится не к x, как в классике, а к этой вероятностной функции. Для случая одной степени свободы Шрёдингер пишет [1]

$$-\frac{h^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V(x)\,\psi = ih\,\frac{\partial\psi}{\partial t}, \tag{4}$$

где $V(x)$ — потенциальная энергия, как она выражается в классике; например, для осциллятора $V(x) = \frac{kx^2}{2}$. «Рецепт» следующий: зная $\psi(x,\ 0)$ при $t = 0$, мы решаем (4) и находим $\psi(x,\ t)$. Искомая вероятность будет

$$\Phi(x,\ t) = \psi(x,\ t)\,\psi^*(x,\ t).$$

Таково изменение распределений со временем.

---

[1] Здесь h — планковская постоянная, деленная на 2π.

Естественно напрашиваются следующие вопросы:

a) как пришли к уравнению (4), как его «вывели»?

b) как писать это уравнение в различных случаях?

Ответ на второй вопрос чрезвычайно прост: берите классическую гамильтонову функцию [для одной степени свободы это $H = \frac{p^2}{2m} + V(x)$] и заменяйте в ней p и x *операторами*: $x \to x$, $p \to -ih\frac{\partial}{\partial x}$; тогда получится уравнение Шрёдингера

$$H\psi = ih\frac{\partial\psi}{\partial t}. \qquad (5)$$

Что касается первого вопроса, т. е. вопроса о том, откуда же взят изложенный рецепт, то надо сказать прямо: *вывода* здесь нет и не может быть. Вывести — это значит получить логическим путем из других законов. Поэтому новые вещи *не выводятся*. Уравнение Шрёдингера *угадано*. Все величие и вся мощь нового аппарата — в его разрыве с классикой. Волновая механика утверждает, что законы микромира принципиально отличны от классических и поэтому их нельзя вывести из классики. Классические законы в микромире просто неверны, и при переходе к новым законам логический скачок должен быть. И он есть. Конечно, можно сказать, что мы не удовлетворены сухим *рецептом*. Эту неудовлетворенность можно устранить, проследив, как *пришли* исторически к (4), но именно как пришли, а не вывели.

Сейчас нам кажется невероятным, как могли долгое время думать, что внутри атомов верны классические уравнения движения. Это, конечно, предрассудок. Классические уравнения взяты из макроскопического опыта, уравнения (2) припасованы к движениям больших тел, больших масс. Отсюда очень далеко до гипотезы, что они же справедливы и в микромире. Это убеждение — предрассудок. Но другое должно иметь место: новые уравнения при переходе к макроскопическим телам должны давать классические уравнения. Этого мы вправе от них требовать, и это требование играло роль при их установлении. Однако простая связь квантов и классики была найдена не попутно с установлением новых уравнений, а лишь впоследствии. В особенно простой форме это сделал Эренфест.

Конечно, и с рецептом дело обстоит сложнее. Мы рассмотрели только координату, но состояние однозначно определено только тогда, когда мы знаем еще и скорость или импульс. Значит, наряду с распределением в пространстве надо еще уметь сказать, что дает ψ для распределения по импульсам, для определения вероятности $G(p)\,dp$ того, что импульс лежит в интервале $(p, p + dp)$. Ответ — вполне определенный и очень общий, но я остановлюсь на этом

коротко, так как сегодня нам это не понадобится. Если мы имеем какой-либо оператор L, то, как известно, его собственными значениями называются те значения параметра λ, при которых уравнение

$$L\varphi = \lambda\varphi$$

имеет решения (эти решения — собственные функции). Для импульса

$$-ih\frac{\partial\varphi}{\partial x} = p\varphi, \quad \text{откуда} \quad \varphi = \text{const}\cdot e^{\frac{ipx}{h}}.$$

Рецепт отыскания $G(p)$ таков: найдите собственные функции интересующей нас величины (p), разложите ψ в ряд (интеграл) по этим собственным функциям φ, и тогда квадраты модулей коэффициентов разложения и будут $G(p)$, соответствующие найденным значениям p.

Итак, и $\Phi(x)$ и $G(p)$ получаются с помощью одной и той же функции ψ, и это влечет за собой наличие связи между ними. Правда, эта связь, получающаяся чисто математически, еще очень широка, но все же она есть, и это — *принципиальное отличие от классики*. Физически нам важно $\psi\psi^*$, но всякое ψ, умноженное на $e^{if(x)}$, даст, очевидно, то же самое $\Phi(x)$. Разложение же ψ по φ будет различно, и $G(p)$ получится иным для ψ и для $\psi e^{if(x)}$. Таким образом, Φ и G связаны, хотя и не очень крепко, но связаны: чем у́же Φ, тем шире G, и наоборот. Количественные оценки ширины обоих распределений соотносятся так, что

$$\Delta x\cdot\Delta p \geqslant h. \tag{6}$$

Это знаменитое соотношение неопределенностей Гейзенберга. То, что классические x и p представляют собой предельные случаи распределений Φ и G, было бы тривиально, если бы не соотношение (6). Оно утверждает, что предельно острые распределения Φ и G *одновременно* невозможны.

Вернемся к проблеме дисперсии. Как мы совместим теперь (1) и (4)? В классике надо было найти $\mathbf{r}$ в зависимости от E в *каждой точке*, причем мы считали — и это, конечно, остается в силе и в квантах,— что в пределах атома, в силу его малости, E не меняется от точки к точке. Между тем в уравнении (4), в отличие от классического уравнения (2), вообще нет понятия об $\mathbf{r}$ в данной точке. Согласно волновой механике, под $\mathbf{r}$ надо понимать *статистическое* среднее.

Вообще средним значением называется математическое ожидание

$$\bar{x} = \sum xW(x),$$

или в нашем случае

$$\bar{x} = \int_{-\infty}^{+\infty} x\psi\psi^* \, dx. \tag{7}$$

(7) можно найти для каждого электрона, и мы получим тогда интересующую нас зависимость: $\bar{x}$ как фунцию t, но не x. Квантовая теория дисперсии постулирует, что надо найти связь $\bar{x}$ с E и ее подставить в (1). Мы и здесь получим, что $\bar{x}$ пропорционален E, и коэффициент пропорциональности войдет в показатель преломления уже известным образом.

Почему в $\bar{x}$ войдет E? Потому, что в (4) теперь, при наличии поля волны, войдет в качестве потенциальной энергии не внутренняя энергия $V(x)$, а

$$V(x) - eEx. \tag{8}$$

Надо решить теперь (4) с (8). Это не просто, и решение находят приближенно. И вот оказывается, что если проделать это *для осциллятора*, то в первом приближении нет никаких отличий от классики. Для кулонова поля отличия есть, в результаты входит h, что сразу же указывает на недостижимость этих результатов классическим путем. Получаются новые эффекты (например, отрицательная дисперсия), но и в общем случае дисперсионные формулы очень похожи на классические: те же резонасные множители $\frac{1}{\omega_0^2 - \omega^2}$ и т. д. В деталях (конечно, важных деталях) есть отличия.

Итак, надо решать уравнение (4), проделать все выкладки, и в результате получится почти классический результат, а для осциллятора (по крайней мере в первом приближении) — в точности классический. Возникает вопрос: что для осциллятора совпадение случайно или принципиально?

Для совсем других целей Эренфест установил теорему, к которой мы теперь и перейдем. Вопрос заключается в следующем: можно ли как-нибудь просто и наглядно перейти от квантов к классике? Теорема Эренфеста говорит о том, как и когда это можно сделать. Возьмем (7), и посмотрим, какому уравнению удовлетворяет $\bar{x}$. Ведь физически нас интересует именно $\bar{x}$, а вовсе не x, ψ и т. д. Может быть, $\bar{x}$ удовлетворяет более простому уравнению?

Перепишем (4) для ψ^*

$$-\frac{h^2}{2m}\frac{\partial^2\psi^*}{\partial x^2} + V(x)\psi^* = -ih\frac{\partial\psi^*}{\partial t}. \tag{4'}$$

Умножим (4) на ψ^*, а (4′) на $-\psi$ и сложим. Получаем

$$\frac{\partial\psi\psi^*}{\partial t} = \frac{ih}{2m}\left(\psi^*\frac{\partial^2\psi}{\partial x^2} - \psi\frac{\partial^2\psi^*}{\partial x^2}\right) = \frac{ih}{2m}\frac{\partial}{\partial x}\left(\psi^*\frac{\partial\psi}{\partial x} - \psi\frac{\partial\psi^*}{\partial x}\right).$$

Вычислим теперь $\dot{\bar{x}}$

$$\frac{d\bar{x}}{dt}=\int\limits_{-\infty}^{+\infty} x\,\frac{\partial\psi\psi^*}{\partial t}\,dx=\frac{ih}{2m}\int\limits_{-\infty}^{+\infty} x\,\frac{\partial}{\partial x}\left(\psi^*\,\frac{\partial\psi}{\partial x}-\psi\,\frac{\partial\psi^*}{\partial x}\right)dx=$$

$$=\frac{ih}{2m}\left[x\left(\psi^*\,\frac{\partial\psi}{\partial x}-\psi\,\frac{\partial\psi^*}{\partial x}\right)\right]_{-\infty}^{+\infty}-\frac{ih}{2m}\int\limits_{-\infty}^{+\infty}\left(\psi^*\,\frac{\partial\psi}{\partial x}-\psi\frac{\partial\psi^*}{\partial x}\right)dx.$$

Все функции предполагаются такими, что в $\pm\infty$ они достаточно быстро стремятся к нулю. Проинтегрированная часть поэтому пропадает, и мы получаем

$$\dot{\bar{x}}=-\frac{ih}{2m}\int\limits_{-\infty}^{+\infty}\left(\psi^*\,\frac{\partial\psi}{\partial x}-\psi\,\frac{\partial\psi^*}{\partial x}\right)dx.$$

Если таким же путем вычислить $\ddot{\bar{x}}$, то получается *замечательный* результат

$$m\frac{d^2\bar{x}}{dt^2}=-\int\limits_{-\infty}^{+\infty}\frac{\partial V}{\partial x}\,\psi\psi^* dx=-\overline{\frac{\partial V}{\partial x}}\,. \tag{9}$$

Таким образом, ускорение удовлетворяет простому уравнению, но это уравнение ничего не дает: чтобы вычислить интеграл, надо знать ψ, т. е. решить (4). Эренфест использовал (9) так: пусть $\psi\psi^*$ — очень узкое распределение и на всем интервале Δx, где $\psi\psi^*$ имеет заметную величину, $\frac{\partial V}{\partial x}$ почти не меняется. Если ширина пакета Δx не меняется или достаточно мало меняется со временем, то в течение известного промежутка времени можно заменять (9) уравнением, в котором $\frac{\partial V}{\partial x}$ вынесена из-под интеграла. Так как $\int\limits_{-\infty}^{+\infty}\psi\psi^* dx=1$, то

$$m\frac{d^2\bar{x}}{dt}=-\frac{\partial V}{\partial x}\,, \tag{10}$$

т. е. ньютоново уравнение. Оно справедливо для больших тел, больших масс, когда расплывание пакета ничтожно. Таким образом, ньютонова механика есть предельный случай, причем $\bar{x}$ — это ньютоново x.

Эренфест приводит следующие примеры:

1) $m=1$ *г*, ширина пакета $\Delta x=10^{-3}$ *см*. Ширина пакета удвоится через время $T=10^{21}$ *сек*. В течение этого времени можно пользоваться (10);

2) $m = 10^{-12}$ *г*, $\Delta x = 10^{-4}$ *см* (броуновская частица). И здесь еще $T = 10^7$ *сек*;

3) $m = 1{,}7 \cdot 10^{-24}$ *г*, $\Delta x = 10^{-8}$ *см* (электрон в пределах атома), $T = 10^{-13}$ *сек*. Отсюда ясно, что для электрона в атоме ни о каком применении классики не может быть и речи. Там $\frac{\partial V}{\partial x}$ отнюдь нельзя считать постоянной всюду, где $\psi\psi^* \neq 0$. Значит, пользоваться теоремой Эренфеста для дисперсии преждевременно. (9) верно, но (10) нет, и мы опять стоим перед необходимостью проведения полного расчета.

Но возьмем осциллятор. Здесь потенциальная энергия будет $\frac{kx^2}{2} - eEx$, и, следовательно, (9) принимает вид

$$m\frac{d^2\bar{x}}{dt^2} = -k\int\limits_{-\infty}^{+\infty} x\psi\psi^* dx + eE\int\limits_{-\infty}^{+\infty} \psi\psi^* dx,$$

т. е.

$$m\ddot{\bar{x}} + k\bar{x} = eE. \qquad (11)$$

Значит, в этом *частном* случае для среднего значения *совершенно строго* верна классика со всеми вытекающими из (11) последствиями. Конечно, это связано со специальным видом $V(x)$: только для квадратичной и линейной форм и постоянной это так, причем то же самое справедливо и для пространственного и даже анизотропного осциллятора.

Как это ни просто, но это проясняет связь между классической и квантовой теорией дисперсии. Здесь имеются два различия — моделей и принципов. Так вот для модели осциллятора различие в принципах *применительно к вопросам дисперсии* несущественно (в других вопросах, как, например, запас энергии и т. д., разница есть).

При наличии магнитного поля $\mathfrak{H}$ меняется сам вид уравнения Шрёдингера, ибо, как известно, в магнитном поле уже нельзя говорить о потенциальной энергии $V(x)$. Известно, однако, что и в этом случае можно подыскать обобщенную гамильтонову функцию, вводя в импульсы p векторный потенциал. Квантовики берут эту обобщенную классическую гамильтонову функцию и ее переводят на язык операторов. Это и дает обобщенное уравнение Шрёдингера. Вместо (9) тогда получается

$$m\frac{d^2\mathbf{r}}{dt^2} = -\int\limits_{-\infty}^{+\infty} \nabla V \cdot \psi\psi^* d\tau - \frac{e}{c}\,[\mathbf{r}, \mathfrak{H}],$$

т. е. и здесь для осциллятора получается классическое уравнение движения.

Итак, в случае осциллятора все вопросы дисперсии получают в квантовой теории тот же ответ, что и в классике. Мы видим, что важны не кванты, а модель, и, к сожалению, не случайно, что это все получается только для осциллятора. Осциллятор — единственная система, где частота не зависит от амплитуды, а, грубо говоря, квантуется ведь именно амплитуда. Всюду, где частота зависит от амплитуды (нелинейные системы), неизбежно — и нельзя ожидать другого — классика и кванты не дадут одного и того же.

В заключение я хочу указать на то, что сведение к классике может иметь и более серьезное значение при исследовании принципиально новых вопросов, так как оно сильно все упрощает. Я укажу для примера на вопрос о запаздывании явления Фарадея, рассмотренный М. А. Дивильковским [1], вопросы, касающиеся сдвига и уширения спектральных линий, наконец, вопросы излучения. У Дирака они рассмотрены приближенно и не строго — есть расходящиеся ряды и т. д. Я думаю, что это не от существа дела, как, например, в вопросе о бесконечной энергии, где проявляются недостатки теории, а от плохих методов приближения при вычислениях.

[1] [ЖЭТФ, 7, вып. 5, 650, 1937.]

СЕМИНАР ПО ДИСПЕРСИИ И АБСОРБЦИИ

(1936—1937 гг.)

ВВОДНАЯ ЛЕКЦИЯ

(14.XI 1936 г.)[1]

Темой нашего семинара является дисперсия. Я хотел бы рассмотреть сегодня некоторые общие вопросы, очертить сначала, хотя бы начерно, круг проблем, которые нас будут интересовать, и указать на объединяющие их точки зрения.

Этимологически «дисперсия» означает рассеяние, разброс. Слово «дисперсия» укоренилось поэтому для обозначения самых различных понятий. В статистике дисперсией называют среднюю квадратичную ошибку, меру разброса значений случайной величины от среднего. В оптике говорят о дисперсии осей в кристаллах, понимая под этим различные направления осей для разных длин волн. Но нас будет интересовать круг явлений, берущий свое начало в известных опытах Ньютона над разбросом цветов в призме.

Уже вскоре после Ньютона искали причину этого явления в различии скоростей распространения лучей разной окраски. Сам Ньютон держался только факта различной преломляемости таких лучей, но это скоро было сведено к разной скорости волн или корпускул. Для правильного объяснения корпускулярная теория света должна была предположить при этом, что скорость в среде больше, чем в воздухе; волновая теория — наоборот. Это противоречие возникло очень рано, и вслед за тем был объявлен experimentum crucis — непосредственное измерение скорости света. Думали, что измерения Физо и Фуко подтверждают волновую теорию, но сегодня мы знаем, что подобный решающий опыт, быть может правильный в деталях, принципиально неверен. Какая скорость непосредственно измеряется в этих опытах? Что вообще пони-

[1] [Из материалов семинара по дисперсии и абсорбции удалось воспроизвести вводную лекцию Л. И. Мандельштама и его выступление в связи с докладом М. А. Леонтовича о распространении волнового фронта. При этом целесообразно было поместить также конспект самого доклада М. А. Леонтовича. Материал обработан С. М. Рытовым по собственным записками.]

мать под скоростью света в среде, где различные волны распространяются с различной скоростью? Эти вопросы касаются самой постановки разных опытов (измерение показателя преломления, определение скорости методами Физо и Фуко, измерения коэффициента увлечения, аберрации и т. д.) и открывают огромный круг проблем.

Понятие скорости при волнообразном распространении само по себе совсем не простая вещь, и его следует уяснить прежде всего. Исторически понятие скорости возникло в механике из движения точки. Чтобы охарактеризовать состояние движения точки, чтобы охватить всю сложность или, если хотите, простоту этого явления, мало дать координату точки. Мы говорим, сверх того, еще о скорости точки, *определяя* это понятие как

$$\lim_{\Delta t\to 0}\frac{\Delta l}{\Delta t},$$

и давая различные рецепты для измерения Δl и Δt. Это не единственный способ определения скорости. Можно было бы (и это было бы естественно с точки зрения физических понятий) непосредственно ввести *эталон* скорости, например скорость определенного ядра из определенной пушки, и дать такой рецепт измерения: скорость, равная 2, есть скорость ядра, выпущенного из такой же самой второй пушки, движущейся вместе с ядром из первой пушки. Теория относительности позволяет утверждать, что в области малых скоростей оба определения совпадают, но совершенно расходятся для скоростей, приближающихся к c,— скорости света в пустоте. Второе определение используется больше всего



Рис. 1

английскими авторами и обозначается ими, в отличие от скорости, термином «быстрота» (rapidity).

Обычное определение удобно и тем, что оно тесно примыкает к математике: для решения уравнений второго порядка необходимо задание не только x_0, но и $\dot{x}_0$, т. е., говоря физически, не только начальной координаты, но и начальной скорости в обычном ее понимании.

Что существенно для *всякого* определения — это возможность отождествить точку в моменты t и $t+\Delta t$, с одной стороны, и воз-

можность измерить координату, *не нарушая* самого движения,— с другой. Развитие физики показало, что выполнение обеих этих предпосылок отнюдь не является самоочевидным, как это считала классическая физика. Но нас сейчас интересует другое.

Говорят так: перенесем это понятие скорости на (любые) волны. Оправдан ли такой перенос? Я думаю, что нет, и вот почему. В волновом движении мы имеем *перенос состояния*. Чтобы судить о скорости переноса, надо уметь отождествлять поле (состояние) в одном и в другом месте, в один и в другой момент. Но для этого нет никаких данных: мы не можем сказать, что E' это есть E, переместившееся из A в B к моменту t'. Здесь нет отождествления, и тем самым — по крайней мере в общем случае распространения состояний — понятие скорости становится неопределенным. Его здесь просто нет, и я думаю, что оно здесь и не нужно. Но проблемы, которые будут интересовать нас, несколько у́же, и для них дело обстоит проще, хотя анализ вводимых понятий, конечно и здесь совершенно необходим.

Нас интересует распространение волн в среде. Ее свойства, если речь идет о непрерывной среде, о *континууме*, мы выражаем дифференциальными уравнениями в частных производных. Будем рассматривать для простоты *плоскую* задачу, т. е. такую, в которой пространственная зависимость входит только через одну координату. Далее, нас главным образом интересуют те случаи, когда справедлив принцип суперпозиции и соответственно дифференциальные уравнения являются линейными. Это отнюдь не значит, что практически всегда идет речь о линейных уравнениях. Совсем недавно практика поставила вопрос о так называемом люксембургско-горьковском эффекте. Приемник B, настроенный на



Рис. 2

волну станции A, воспринимает модуляцию мощной промежуточной станции. По-видимому, здесь дело именно в отсутствии суперпозиции волн в ионосфере, в результате чего эффект действия двух станций не равен сумме эффектов от каждой из них. Таким образом, имеются проблемы, в которых основной интерес лежит как раз в нарушении принципа суперпозиции, но достаточно большой круг вопросов описывается линейными уравнениями.

Наконец, мы часто имеем дело с однородными и постоянными средами, т. е. такими, свойства которых не зависят от точки и от

времени. Так как свойства среды входят в дифференциальное уравнение через его коэффициенты, то в таком случае мы получаем дифференциальные уравнения с постоянными (и большей частью действительными) коэффициентами.

Итак, мы имеем одно или несколько линейных дифференциальных уравнений с постоянными коэффициентами. Такие уравнения допускают частные решения вида

$$e^{i(kx-\omega t)}, \quad \text{где } k = k(\omega),$$

т. е. в данной среде возможны такого рода процессы. Мы имеем здесь функцию $\left(x - \frac{\omega}{k} t\right)$, а не просто от x и от t порознь. Для таких процессов можно говорить о скорости

$$v = \frac{\omega}{k}, \quad \text{где } k = \frac{2\pi}{\lambda}.$$

Таким образом, для определенного типа возмущений, для синусоидальных волн (бесконечно протяженных по x и t) может быть введена скорость как некоторое вполне определенное понятие. Но нетрудно показать, что именно этот весьма частный тип процессов нас меньше всего интересует. Если движение не синусоидально, то v сохраняет смысл в *отсутствие дисперсии*. Мы не можем, в частности, обрывать синусоиду, ибо такой обрыв находится в противоречии с самим определением. Можно говорить об ограниченной синусоиде, но по отношению к ней нельзя делать никаких выводов из сказанного выше о синусоиде бесконечной, нельзя отсюда извлекать какие-либо следствия. Между тем это делали, и это часто приводило к недоразумениям.

Что же нашел в своем опыте Фуко? Он искал скорость распространения света, именно ту скорость, которая входит в закон преломления. Между тем в самом принципе опытов Физо и Фуко заложена несинусоидальность волн (прерывание света). Следовательно, в этих опытах измеряется *не* фазовая скорость v и тем самым experimentum crucis не получается. Для скорости света в воде как будто бы все вышло правильно, но для сероуглерода Майкельсон нашел из измерения скорости света по методу Физо значение показателя преломления $n = 1{,}77$, в то время как из преломления $n = 1{,}64$. Это было загадкой, пока Релей не указал, что во всех прямых измерениях скорости распространения определяется *групповая*, а не фазовая скорость.

Известно далее, что принцип относительности по соображениям, связанным с причинностью, не допускает для любого тела или сигнала скорости, большей скорости света в вакууме c. Но фазовая скорость $v = \frac{\omega}{k}$ сплошь и рядом бывает больше c. В этом

тоже видели противоречие, упуская из виду, что v — скорость бесконечной синусоидальной волны, т. е. не скорость сигнала, ибо такой волной ничего нельзя передать. Всякий сигнал по необходимости означает нарушение синусоидальности. Но тогда и скорость уже не будет v. Этот вопрос тоже привлекал много внимания, хотя возник он просто из некритического переноса механических понятий в область распространения волн.

Теперь такие вопросы более или менее ясны, но ряд понятий здесь надо себе уяснить: скорость группы, скорость энергии, скорость фронта и т. д.

Можно ли подойти к дисперсии, вообще не вводя понятия скорости (помимо v)? Это, по-моему, нецелесообразно. В целом ряде случаев условия не являются чрезмерно общими, и тогда представляется возможным и разумным — конечно, при наличии критического подхода — ввести сильно облегчающее понятие скорости. Как я уже сказал, хотя многое здесь и известно, все же многое нуждается в уяснении. В каких пределах применимо понятие групповой скорости? Каково ее соотношение с волновой скоростью, со скоростью энергии, скоростью фронта и т. п.? В волновой механике эти вопросы — основа всего. Поэтому в качестве первого пункта программы нашего семинара мы поставим вопросы, касающиеся распространения волн. Таким образом, первый цикл, охватывающий ряд общих положений, относящихся к различным случаям кривых дисперсии, составит *кинематика распространения волн в диспергирующих средах*.

Далее возникает вопрос о том, как свойства среды определяют тот или иной конкретный вид дисперсии. Свойства среды определяют коэффициенты дифференциальных уравнений, а эти коэффициенты, в свою очередь, определяют закон дисперсии. Можно ставить и обратный вопрос — о том, как из дисперсии сделать заключения о свойствах среды. Поэтому второй цикл вопросов будет состоять в *разборе некоторых типичных и важных сред*.

Мы встречаемся с диспергирующими средами в самых различных областях физики — и в оптике, и в акустике, и в радио. Вопросов очень много, и необходимо поэтому выделить наиболее важные. В числе последних прежде всего, как мне кажется, надо назвать *изучение сред с помощью оптической дисперсии*. Это мы и поставим на первом месте из вопросов второго цикла.

Проследим коротко, как развивалась теория оптической дисперсии, какие здесь были извилины в ходе идей, что отпало и что сохранилось до нашего времени.

Начало — это классические опыты Ньютона, которым он сам придавал огромное значение (в письме секретарю Королевского Общества он говорит, что «я излагаю здесь исключительно важные опыты») и которые были поставлены с непревзойденным экспери-

ментальным искусством. Но о свойствах среды сам Ньютон, по-видимому, не ставил вопроса и, более того,— существенно затормозил его развитие. Он неразрывно связал дисперсию с преломлением и считал поэтому, что дисперсия принципиально неустранима. Это практически чрезвычайно существенный момент, ибо совершенные телескопы, основанные на преломлении (рефракторы), с этой точки зрения невозможны. Так думали в течение почти сотни лет после Ньютона. Сам Ньютон придумал отражательный телескоп-рефлектор.

Эйлер, намного опередивший свое время, первый ясно представил себе контуры волновой теории света вплоть даже до утверждений, что цвет связан с длиной волны λ и что красные лучи имеют бóльшую λ, а фиолетовые — меньшую. Он говорил: глаз (почти) ахроматичен, но он состоит из нескольких сред. Значит, система из нескольких сред может быть лишена дисперсии, оставаясь преломляющей. Эйлеру усиленно возражал Доллонд, который в 1759 г. сам построил из флинта и крона первый ахроматический объектив.

Однако первые серьезные взгляды в этом вопросе были высказаны Френелем и Коши. Последний дал дисперсионную формулу

$$v = a + \frac{\alpha}{\lambda^2} + \frac{\beta}{\lambda^4} + \ldots$$

Из каких соображений они исходили? Возможный ход идей рисуется следующим образом. Если эфир — *непрерывная* среда с плотностью ρ и упругостью E, то дисперсии нет. Следовательно, дисперсия обусловлена влиянием *частиц*. v зависит от λ потому, что внутренние свойства среды связаны с λ или ω (соответственно пространственные или временны́е свойства). Если λ гораздо больше пространственной зернистости среды (или ω гораздо меньше каких-то внутренних собственных частот), то дисперсии ожидать нельзя, если же λ (или ω) сравнима с собственными пространственными (или временны́ми) масштабами, то дисперсия будет иметь место.

Френель и Коши видели только влияние пространственной неоднородности. Они считали, что сами молекулы не участвуют в распространении света, но они уплотняют эфир, делают его зернистым. Из этих представлений Коши и получил свою формулу. В этом же направлении шли затем Кеттелер и Гельмгольц. Эти первые схемы пригодны только для объяснения нормальной дисперсии, для объяснения же открытой Леру в 1860 г. аномальной дисперсии этот путь отпадал. Идея о влиянии временны́х свойств среды еще не созрела к тому времени, и опыты Леру остались в тени. Однако позднее и теория пришла к пониманию роли соотношения между ω и собственной частотой ω_0 частиц среды, к представлению

о *резонаторах*, вкрапленных в упругий эфир. В работах Кеттелера, Гельмгольца, Зеельмейера теория была доведена почти до полного завершения, почти до современного состояния. И последующие опыты, «открытие» аномальной дисперсии в фуксине (главным образом работы Кундта), произвели фурор, переворот. Очень скоро увидели, что если приписать резонаторам трение, то получается также и объяснение абсорбции. В таком виде теория, существенно отличная в основных чертах от теории Френеля — Коши, перешла затем в электромагнитную теорию света.

Любопытно отметить, что вопрос о распространении света в *упругой* среде с вкрапленными в нее разонаторами был решен Релеем как экзаменационный вопрос, поставленный на экзамене Максвеллом (1869 г.). Таково же, между прочим, происхождение и теоремы Стокса!

В классической электронной теории общая теория дисперсии приобрела вполне развернутый вид и праздновала здесь свои триумфы — объяснение явления Зеемана, явления Фарадея, резонансной флуоресценции и др. Мы имеем теперь квантовую теорию дисперсии и абсорбции. В отношении результатов она внесла сравнительно немного нового. Это новое она вносит двояко: во-первых, она показывает, что непригодна модель осциллятора; во-вторых, она изменяет весь подход к явлениям. Принципиально новое следствие квантовой теории — так называемая отрицательная дисперсия, экспериментально подтвержденная Ладенбургом. Мы еще вернемся к вопросу о том, является ли отрицательная дисперсия результатом изменения модели или же изменения теории, т. е. можно ли с другой моделью получить отрицательную дисперсию классически. Наконец, к этому же кругу вопросов относятся экспериментальные методы исследования дисперсии и абсорбции.

Другая группа вопросов второго цикла касается классификации типов дисперсии. В классической теории оптической дисперсии главным моментом является инерция электронов. Именно она определяет характер явлений, наступающих при приближении ω к ω_0. Это *инерционный* тип дисперсии, для которого характерно, что дисперсия и абсорбция, вообще говоря, связанные друг с другом, определяются *разными* постоянными — массой и коэффициентом трения. Инерционная дисперсия имеет место и в отсутствие трения, ее можно назвать поэтому консервативной дисперсией. Заслуга общей теории оптической дисперсии состоит в частности, в том, что наряду с инерционной дисперсией она охватывает также абсорбционную. Но имеются такие области, где основное — абсорбция, а не дисперсия. Так обстоит дело, например в средах со свободными электронами (металлы, ионосфера). Конечно, инерция будет давать известную дисперсию, но главное — это абсорбция, обусловленная большим числом соударений.

Необходимость создания теории среды со свободными зарядами возникла в радиотехнике в связи с распространением электромагнитных волн в ионосфере. Уже в самом начале развития радио, в 900-х годах, заметили, что дальности радиоприема получаются много больше ожидаемых. Связь с Америкой считали невозможной, и нужна была известная смелость для того, чтобы сказать: теория теорией, а посмотрим, что получится? Именно так поступил Маркони. Он попробовал — и получилось. Тогда Хивисайд, аргументируя зависимостью дальности от времени года и дня, от широты и т. д., выдвинул гипотезу об ионизации верхних слоев атмосферы и о наличии у этих слоев соответствующей проводимости (одновременно эту гипотезу высказал Кеннели). К тому же пришли еще в 80-х годах Шустер и Стюарт в поисках объяснения для большой энергии магнитных бурь, превосходящей энергию, доставляемую Солнцем.

Конечно, было ясно, что магнитные бури связаны с Солнцем (влияние пятен), но энергия их слишком велика. Шустер и Стюарт предположили, что Солнце ионизирует верхние слои атмосферы, а движение этих слоев, дающее бури, черпает свою энергию уже из вращения Земли [1]. В последнее время, особенно благодаря переходу к коротким волнам, был открыт ряд новых фактов (фединги, ночные и дневные волны), тоже заставляющих глубоко интересоваться верхними слоями атмосферы.

Позднее Лармор дал теорию дисперсии в среде со свободными электронами [2]. Плотности в ионосфере малы, свободный пробег может достичь около 1 *м* и более, так что электрон успевает совершить много вынужденных колебаний, давая свою инерционную дисперсию, прежде чем ударится и отдаст полученную энергию (абсорбция). Лармор дал первый расчет (с учетом абсорбции). Его формула для скорости волны, если пренебречь абсорбцией, имеет следующий вид:

$$v^2 = \frac{c^2}{1 - \dfrac{Ne^2}{\pi m \nu^2}},$$

где e и m — заряд и масса электрона, N — число электронов в 1 $см^3$. Таким образом, v всегда больше c и с увеличением частоты ν стремится к c.

В той области частот ν, где знаменатель близок к нулю, уже

---

[1] [Л. И. Мандельштам имеет здесь в виду теорию так называемого атмосферного динамо, согласно которой магнитное поле токов, возникающих в ионизованном слое благодаря перемещению зарядов, вызывает вариации земного магнитного поля. См. *B. B. Stewart*. Encycl. Brit., 9, 181, 1878; *A. Schuster*. Phil. Trans. Roy. Soc., 180, 467, 1889.]

[2] [*J. Larmor*. Phil. Mag., 48, 1025, 1924.]

нельзя пренебрегать соударениями. Наконец, при дальнейшем удлинении волны мы переходим в область, где $v^2 < 0$. Что это значит? Грубо говоря, это значит, что волна полностью отражается от такой среды. Переход к такому полному отражению сравнительно очень резок, и это дает один из хороших методов исследования ионосферы. Укажем для иллюстрации, какого рода здесь возникают вопросы, решение которых имеет жизненный интерес как для радио, так и для геофизики.

Во-первых, на какой высоте находится ионизованный слой? Хивисайд считал, что около 100 *км*, и это в общем подтвердилось. Метод определения весьма прост: передатчик посылает сигнал — короткий цуг, который распространяется, конечно, с групповой скоростью. Стоящий неподалеку приемник принимает этот сигнал один раз непосредственно от передатчика (земной луч), а затем после отражения от ионосферы (эхо). Запоздание позволяет определить высоту слоя.

Эпльтон нашел, далее, что при укорочении волны при частоте в $3{,}5 \cdot 10^6$ циклов (в его опыте) высота слоя скачком увеличилась примерно до 200—250 *км*. Он объяснил это явление наличием второго, более высокого и более плотного слоя: при переходе ν через $3{,}5 \cdot 10^6$ циклов нижний слой перестал отражать, но отражал верхний. Так был открыт второй слой, который правильно было бы назвать слоем Эпльтона. Таким образом, распространение сигнала позволяет найти высоту слоя, а дисперсия дает способ узнать N по граничной (критической) частоте

$$N = \frac{\pi m}{e^2} \nu_{\text{гр}}^2.$$

Первый слой назван слоем E, второй — слоем F (как оказалось, он в некоторых случаях двоится, так что различают слои F_1 и F_2).

Далее возникает вопрос об ионизующем факторе. По целому ряду соображений (потенциал ионизации, абсорбция солнечного света и т. д.) мы можем с уверенностью утверждать, что ионизация вызывается ультрафиолетовыми лучами. Во всяком случае, это главный ионизующий агент. Но уже в слое F активная часть ультрафиолета поглощается, как предполагал Чепмэн, настолько, что остатка для наблюдаемой ионизации нижнего слоя не хватает [1]. Приписать дополнительную ионизацию только солнечным электронам не представлялось возможным по соображениям, связанным с распределением заряженных частиц по широте (концентрация у полюсов). Поэтому было высказано также предположение, что ионизация обусловлена еще какими-то нейтральными частицами.

---

[1] [*S. Chapman*. Proc. Roy. Soc., 132, 353, 1931. Согласно современным воззрениям, ультрафиолетовой энергии Солнца достаточно для ионизации всех слоев ионосферы.]

Для выяснения этого вопроса особенно удобным моментом является солнечное затмение. Оно позволяет разделить оба фактора (ультрафиолет и корпускулы), так как корпускулярное затмение наступает примерно на 2 часа раньше оптического. Происходит это следующим образом. Для того чтобы дошли до Земли частицы, которые прошли бы через область *A* (рис. 3), если бы их не заслонила Луна, требуется известное время. Поэтому корпускулярная тень будет скошена. При скорости частиц около 2000 *км/сек* это скашивание составляет 1°. Для световой тени оно — соответственно гораздо большей скорости света — значительно меньше, всего лишь около 20″. Так как во время затмения скорость Луны относительно Солнца примерно на 1 *км/сек* меньше, чем Земли, то и получается, что корпускулярное затмение примерно на 2—2,5 часа опережает световое.

Различие в форме полной световой тени (конус, дающий на поверхности Земли область примерно 150 *км* в поперечнике) и тени корпускулярной (цилиндр диаметром 2000 *км*) получается потому, что каждая точка солнечной поверхности излучает свет по всем направлениям, а частицы, как предполагают,— только нормально к поверхности.

Пока нет уверенности в том, что удалось наблюдать корпускулярное затмение [1]. Достоверно установлено, что изменения в ионосфере сопровождают оптическое затмение и, следовательно, ионизация обусловлена ультрафиолетом.

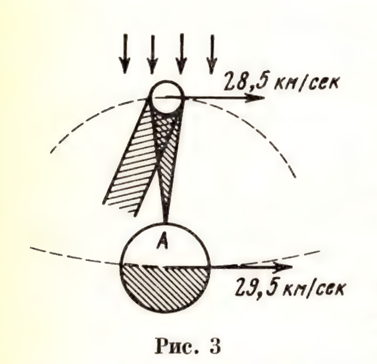


Рис. 3

Во всяком случае, мы видим, что опыты над определением N, над изменением высоты слоев и N в зависимости от различных факторов имеют решающее значение для наших представлений об ионосфере.

Из таких наблюдений оказалось, например, что благодаря магнитному полю Земли в ионосфере получается эффект Фарадея (распространение с различной скоростью двух циркулярно-поляризованных лучей и как результат вращение плоскости поляризации радиоволны).

[1] [Современные сведения о взаимодействии «солнечного ветра» и корпускулярных потоков с магнитосферой Земли исключают прямое попадание частиц в ионосферу. Вместе с тем интенсивные потоки протонов, возникающие при солнечных вспышках, проникают через нейтральные области магнитосферы в авроральную ионосферу, вызывая там ряд явлений (интенсивное рентгеновское излучение, так называемые блекауты и т. п.). См. *W. Hess*. Radiation Belts and Magnetosphere. Blaisdell Publ. Co., Waltham, Massachusetts, Toronto — London, 1968.]

Особо стоит вопрос о так называемых *дальних* эхо, соответствующих высоте слоя в 1,5 млн. *км*. Одна из высказанных гипотез заключалась в том, что дальние эхо обусловлены малостью групповой скорости в ионосфере, но это, по-моему, неверно. Вероятней другая гипотеза, связывающая дальние эхо со штёрмерской теорией северных сияний [1]. Это чрезвычайно интересная гипотеза, согласно которой наши радиосигналы должны выходить далеко за орбиту Луны, но и она встречает ряд затруднений.

Остановимся теперь на двух дальнейших конкретных проблемах второго цикла. Это специальный тип дисперсии в диэлектриках и акустическая дисперсия.

Механизм поляризации диэлектриков в любом электрическом поле тот же, что и при оптической дисперсии,— это смещения связанных электронов, или, точнее, деформация электронных оболочек. Электрон останавливается при действии постоянного поля благодаря тому, что он привязан к своему положению равновесия силой, имеющей характер квазиупругой. Собственные частоты таких осцилляторов весьма велики — это оптические частоты. Волна в 5—6 *м* еще чрезвычайно далека от собственной частоты, она, конечно, вызывает поляризацию, но дисперсия практически будет отсутствовать. Между тем во многих диэлектриках в области коротких волн (3—200 *м*) обнаруживается сильная дисперсия, причем на нее существенно влияет, особенно для жидкостей, температура. С точки зрения инерционной дисперсии это непонятно.

Дебай предположил, что, кроме упругого механизма поляризации, имеется и совершенно иной механизм, связанный с наличием *полярных* молекул. Предположим, что имеется разбавленный раствор (т. е. можно пренебречь взаимодействием молекул растворенного вещества), причем молекулы обладают собственным постоянным электрическим моментом. В отсутствие внешнего поля средний момент равен нулю, так как все направления равновероятны. Поскольку молекулы нейтральны, при наложении поля на каждую будет действовать только пара сил, поворачивающая момент молекулы вдоль по полю. Если бы никаких сил больше не было, то при сколь угодно слабом внешнем поле получилось бы насыщение. Но этого нет. Дебай указывает на антагонистическую силу — удары молекул растворителя, которые препятствуют ориентации и заставляют молекулу совершать вращательное броуновское движение (диффузия моментов). Так как в среднем будет пре-

[1] [Л. И. Мандельштам имеет в виду теорию Штёрмера об образовании под действием магнитного поля Земли частицами, исходящими от Солнца, тороидальной поверхности с осью, совпадающей с магнитной осью Земли. Согласно предположению Штёрмера, происходит отражение радиосигналов от этой поверхности, состоящей из электронов. См. *C. Störmer*. Nature, **122**, 681, 1928.]

валировать направление моментов по полю, то поляризация будет отлична от нуля, а при слабых полях — будет пропорциональна полю. Эта интересная концепция оправдала себя, по-видимому, полностью. Наряду с другими результатами она дает сильную зависимость диэлектрической постоянной от температуры (уменьшение ε с ростом T как в постоянном, так и в переменном поле).

Предположим теперь, что поле меняется настолько быстро, что установление среднего момента не поспевает за ним (броуновская диффузия отстает). Тогда ε будет меньше своего статического значения (дисперсия) и обязательно будет также абсорбция, ибо такие быстрые процессы уже необратимы. Частоты, при которых это происходит, сравнимы со временем релаксации (временем установления), т. е. с обратной величиной k в законе установления e^{-kt}. Это особый тип дисперсии, которую правильно было бы назвать *релаксационной дисперсией*. Изучение такого рода дисперсии и абсорбции дает много сведений об электрических моментах полярных молекул, о свойствах среды, о динамике поворота молекул, об испытываемом ими трении. Изучение этого трения существенно для вопросов тушения флуоресценции, для теории электропроводности электролитов (подвижность ионов), для вязкости (молекулярное и молярное трение) и т. д.

Здесь имеется еще один момент, который на первый взгляд кажется парадоксальным. При упругом типе поляризации энергия слагается из энергии поля $E^2/8\pi$ и энергии растянутых «пружинок», т. е. потенциальной энергии упругих диполей. В сумме с электрической энергией эта упругая энергия дает полную энергию $\varepsilon E^2/8\pi$. Но если речь идет о полярных молекулах, то «пружинок» нет и упругой энергии нет. При тех же значениях ε и E полная энергия не будет равна $\varepsilon E^2/8\pi$ и, как ясно из нашего примера, не может быть этому равна. Здесь часто забывают обстоятельство, которое мы разберем потом подробно. Плотность (полной) энергии выражается как $\varepsilon E^2/8\pi$ только в том случае, если ε не зависит от температуры. Если же зависит, то тогда из термодинамики — в дополнение к модельной теории — следует, что указанное выражение дает только *свободную*, а не всю энергию системы.

Последний класс задач второго цикла — акустическая дисперсия. Акустика считалась классической областью отсутствия дисперсии. Это говорилось во всех учебниках. Даже Релей так думал и привел в подтверждение следующий пример: если бы в воздухе была дисперсия, т. е. звуки разной высоты распространялись с разной скоростью, то на известном расстоянии от оркестра должна была бы получиться какофония. Этого нет, а значит, дисперсии нет. Конечно, знали о том, что ультразвуковые волны с периодом, сравнимым со временем свободного пробега молекулы (10^{-9} *сек*). уже обнаружат некоторые особенности, дадут сильную абсорбцию,

Но соответствующие частоты, порядка 10^8, слишком высоки. При этом не видели возможности возникновения дисперсии и абсорбции при меньших частотах, так как не предполагали наличия более медленных временны́х характеристик среды. Между тем такие характеристики нашлись, и история этого вопроса очень интересна. В 1904 г. Джинс в своем учебнике кинетической теории газов указал на такое явление, дал основную идею и расчет, но в последующих изданиях это место выбросил! [1] Как мы знаем теперь, речь идет о частотах порядка 10^4—10^5. То, что это явление не было тогда интересно и прошло мимо внимания самого Джинса, объясняется уровнем техники. Пьезокварцевой методики тогда еще не существовало.

В 1920 г. к этим вопросам, правда в несколько ином аспекте (применительно к частично диссоциированным газам), вернулся Эйнштейн. За последние 6—7 лет они сделались актуальны. В общих чертах дело заключается в следующем. Скорость распространения звуковой волны — чередующихся сгущений и разрежений — есть $v = \sqrt{\frac{\partial p}{\partial \rho}}$. Она не зависит от частоты. Если процесс изотермический, то $p = \mathrm{const} \cdot \rho$, и тогда $v = \sqrt{\frac{p}{\rho}}$, как нашел Ньютон. Но при звуковых частотах сжатия и разрежения происходят настолько быстро, что правильнее считать процесс адиабатическим. Тогда $v = \sqrt{\gamma \frac{p}{\rho}}$, где $\gamma = \frac{c_p}{c_v}$. Чем больше γ, тем больше скорость. Но чем отличаются c_p и c_v? c_p — теплоемкость при постоянном давлении — равна c_v + работа расширения. Она определяется энергией поступательного движения частиц наряду с другими, внутренними видами энергии — вращательной и колебательной, в то время как c_v, если не учитывать сил сцепления, что возможно, конечно, лишь в случае газов, определяется только энергией поступательного движения. Звуковая волна непосредственно меняет только энергию поступательного движения молекул, но затем эта энергия распределяется и на внутренние степени свободы. При очень высоких частотах последние не успевают возбудиться, и газ будет вести себя так, как если бы у него не было внутренних степеней свободы. Тогда γ и v велики. Происходят колебания быстро или нет, этот вопрос решается сравнением периода волны со временем перехода энергии от внешних степеней свободы к внутренним. Это время перехода и является новой независимой

---

[1] [Джинс имел в виду ротационные степени свободы. Он рассчитал случай, когда молекулы состоят из двух одинаковых соприкасающихся шаров, и показал, что вплоть до частот порядка 10^5 поправка к скорости ничтожно мала. См. *J. H. Jeans*. The dynamical theory of gases. Cambridge, 1904, 277, § 200, стр. 365 и 366.]

временно́й характеристикой среды, раньше не принимавшейся во внимание. Здесь мы имеем *релаксационный* тип дисперсии (если сначала была, скажем, только поступательная энергия, то потом, если систему предоставить самой себе, происходит перераспределение — релаксация). Ясно далее, что здесь получится не только дисперсия, но и абсорбция, так как передача энергии происходит быстро (необратимо). Мы имеем здесь, таким образом, новый способ изучения молекул в газах — способ, открывающий широкие перспективы. В частности, большой интерес представляет изучение влияния примесей, зависимости времени релаксации от температуры и т. п.

Интересную область открывает перенос этих взглядов на жидкости, в которых, по-видимому, тоже обнаружена не только аномальная абсорбция, но и дисперсия. Хотя для жидкостей теория довольно расплывчата, но принципиально явления носят тот же характер, что и в газах. Общая теория такого рода объясняет очень многое и сулит ряд ценных выводов, касающихся и оптики.

Итак, во втором цикле мы рассмотрим следующие конкретные проблемы дисперсии: 1) оптическая дисперсия, 2) ионосфера, 3) дисперсия в полярных диэлектриках и 4) дисперсия в акустике.

РАСПРОСТРАНЕНИЕ ФРОНТА ВОЛНЫ

(Доклад М. А. Леонтовича 14.I 1937 г.)

Мы рассмотрим теперь вопрос о том, с какой скоростью бежит начальная точка некоторого импульса, во всех отношениях произвольного, но такого, что в любой точке пространства до некоторого момента времени вообще нет никакого возмущения. Конечно, вопрос о возможности фактического измерения этой скорости более сложен: импульс может быть или может сделаться очень пологим, так что момент появления возмущения зависит от чувствительности аппарата. Говоря о начальной точке возмущения, мы имеем в виду предельно чувствительный аппарат. Скорость фронта представляет интерес в ряде случаев, особенно в связи с принципом относительности. Последний запрещает скорости, превышающие c. Но существуют области аномальной дисперсии, где фазовая скорость $v > c$, и если бы оказалось, что скорость фронта есть фазовая скорость, то мы получили бы противоречие. Конечно, для мо-

нохроматической волны противоречия нет, так как у нее нет фронта, она не является сигналом. Но для импульса, хотя бы односторонне ограниченного, вопрос остается в полной силе.

Будем по-прежнему рассматривать одномерную задачу, но включим возможность поглощения. Предположим, что волновое поле описывается n линейными однородными уравнениями первого порядка для n функций $\varphi_1, \varphi_2, \ldots, \varphi_n$:

$$\sum_m \left(a_{jm} \frac{\partial \varphi_m}{\partial t} + b_{jm} \frac{\partial \varphi_m}{\partial x} + c_{jm} \varphi_m \right) = 0, \qquad j = 1, 2, \ldots, n. \tag{1}$$

Среду мы будем считать однородной, т. е. все коэффициенты a_{jm}, b_{jm}, c_{jm} постоянны. Разумеется, исходные уравнения могут быть второго и высшего порядков, но, понимая под частью функций φ_m производные от остальных φ, мы всегда можем считать, что задача описывается системой уравнений первого порядка.

Могут быть поставлены два вопроса.

1) Как распространяется монохроматическая волна

$$\varphi_m = a_m e^{i(\omega t - kx)} \qquad ? \tag{2}$$

Решение этого вопроса дает нам закон дисперсии $v = \frac{\omega}{k} = v(k)$.



Рис. 4

Подставив (2) в (1), мы получаем для определения a_m систему однородных уравнений

$$\sum_m (i\omega a_{jm} - ikb_{jm} + c_{jm}) a_m = 0.$$

Условие существования нетривиального решения ($a_m \neq 0$) и дает закон дисперсии

$$\mathrm{Det} \left| a_{jm} v - b_{jm} + \frac{c_{jm}}{ik} \right| = 0, \qquad \therefore v = v(k). \tag{3}$$

2) Какова скорость фронта волны и от чего она зависит? Движение фронта изображается на плоскости (x, t) линией (заранее ясно, что в однородной среде это будет прямая, но это, впрочем, несущественно), по одну сторону которой все $\varphi_m = 0$, а по другую все или хотя бы одна из них отлична от нуля. Теорема, которую мы

хотим доказать, такова: *Если закон дисперсии есть $v = v(k)$, то фронт любой волны идет со скоростью $F = v(\infty)$*, т. е. соответствует скорости бесконечно коротких волн (бесконечно высокой частоте).

Мы воспользуемся следующей математической теоремой: если на отрезке какой-либо кривой на плоскости (x, t) заданы значения всех φ_m, то можно найти единственное решение дифференциальных уравнений (1), обращающееся на отрезке A в заданные значения, причем это возможно лишь при условии, что отрезок A взят *не на характеристике*. Характеристиками называются специальные кривые на плоскости (x, t), обладающие тем свойством, что задание на этих кривых либо вообще не дает решения, либо определяет его неоднозначно (теорема Коши: задание на любой кривой, кроме характеристики, определяет решение однозначно). Нетрудно видеть, что движение фронта должно быть характеристикой. Действительно, мы имеем на линии, изображающей движение фронта (линии Ф), что все $\varphi_m = 0$. Но для таких значений φ_m, заданных на Ф, уравнения (1) заведомо допускают *два* решения. Это, во-первых, рассматриваемый сигнал и, во-вторых, тривиальное решение $\varphi_m \equiv 0$. Следовательно, задание φ_m на линии Ф не определяет решения однозначно, т. е. Ф должно быть характеристикой.

Как же находить характеристики?

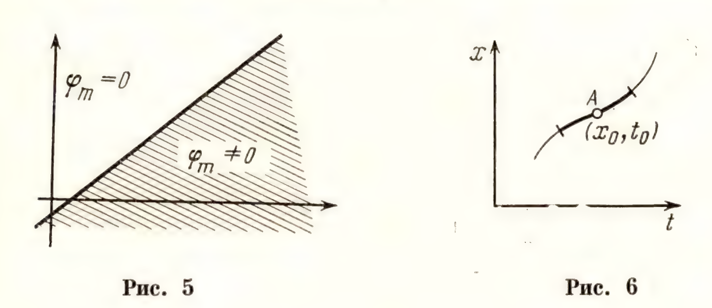


Рис. 5 Рис. 6

Пусть на плоскости (x, t) имеется некоторая кривая, на ней отрезок A и (x_0, t_0) — некоторая точка этого отрезка. Для всего отрезка заданы значения $\varphi_m = \varphi_m^{(0)}$. Индексом 0 мы будем отмечать значения в точке (x_0, t_0).

Будем искать решение в виде рядов

$$\varphi_m = \varphi_{m0} + \left(\frac{\partial \varphi_m}{\partial t}\right)_0 (t - t_0) + \left(\frac{\partial \varphi_m}{\partial x}\right)_0 (x - x_0) + \ldots$$

Мы должны, таким образом, найти коэффициенты этих рядов и прежде всего — первые производные φ_m по x и по t. Так как φ_m

на отрезке A заданы, то, двигаясь по отрезку, мы можем найти полный дифференциал φ_m

$$d\varphi_m^{(0)} = \left(\frac{\partial\varphi_m}{\partial t}\right)_0 dt + \left(\frac{\partial\varphi_m}{\partial x}\right)_0 dx,$$

причем $\frac{dx}{dt}$ — угловой коэффициент касательной к отрезку A. Иначе

$$\left(\frac{\partial\varphi_m}{\partial t}\right)_0 = -\left(\frac{d\varphi_m}{\partial x}\right)_0 \frac{dx}{\partial t} + \frac{\partial\varphi_m^{(0)}}{dt}. \tag{4}$$

Подчеркнем, что $\frac{dx}{dt}$ известно из уравнения кривой, на которой лежит A, а $\frac{d\varphi_m^{(0)}}{dt}$ известно тогда и только тогда, когда мы движемся по отрезку.

Из уравнений (4) надо определить совместно с уравнениями (1), которым φ_m должны удовлетворять всюду, значения $\left(\frac{\partial\varphi_m}{\partial x}\right)_0$ и $\left(\frac{\partial\varphi_m}{\partial t}\right)_0$. Подставив (4) в (1), мы сначала исключим $\left(\frac{\partial\varphi_m}{\partial t}\right)_0$ и получим следующие уравнения для нахождения $\left(\frac{\partial\varphi_m}{\partial x}\right)_0$:

$$\sum_m \left\{\left(-a_{jm}\frac{dx}{dt} + b_{jm}\right)\left(\frac{\partial\varphi_m}{\partial x}\right)_0 + a_{jm}\frac{d\varphi_m^{(0)}}{dt} + c_{jm}\varphi_m^{(0)}\right\} = 0.$$

Эти неоднородные линейные уравнения имеют единственное решение только тогда, когда детерминант отличен от нуля. В этом случае мы однозначно получим $\left(\frac{\partial\varphi_m}{\partial x}\right)_0$, а с помощью (4) определим затем $\left(\frac{\partial\varphi_m}{\partial t}\right)_0$[1]. Все это может не получиться (может вообще не быть решения или же их может быть не одно, а несколько), если детерминант равен нулю, т. е.

$$\mathrm{Det}\left| a_{jm}\frac{dx}{dt} - b_{jm}\right| = 0. \tag{5}$$

Это уравнение n-й степени для $\frac{dx}{dt}$, определяющее n (постоянных)

---

[1] Аналогичным путем затем определяются последовательно высшие производные — коэффициенты последующих членов разложений φ_m. Их вычисление опять сводится к решению системы линейных неоднородных уравнений, причем детерминант системы, определяющей производные n-го порядка, равен n-й степени детерминанта (5).

значений $\frac{dx}{dt}$, среди которых могут быть, конечно, и кратные значения и комплексно сопряженные. Другими словами, (5) определяет, вообще говоря, n различных угловых коэффициентов, n прямых на плоскости (x, t). Если корни (5) комплексно сопряжены, то таким $\frac{dx}{dt}$ действительных направлений на (x, t) не соответствует.

Эти прямые и есть характеристики: задавая значения φ_m на них, мы не можем получить однозначного решения. Так как линия Ф должна быть характеристикой, то *один из корней* (5) *есть скорость фронта*. Сравнивая (5) с (3), мы заключаем, что $F = v(\infty)$, так как при $k \to \infty$ (3) переходит в (5).

Эта теория без существенных изменений может быть обобщена на случай неоднородной среды, когда a, b, c — функции точки (и даже времени). Тогда будет функцией точки как скорость монохроматической волны, так и скорость фронта (т. е. линия Ф будет кривая, а не прямая).

Если мы рассматриваем не одномерную задачу, то имеется еще вопрос о *форме* фронта. Этот вопрос решается чрезвычайно просто: скорость фронта соответствует $\lambda \to 0$, т. е. *переходу к геометрической оптике*. Поэтому все, что касается фронта (форма, момент прихода и т. д.), может быть найдено по предельным законам, законам геометрической оптики — принципу Гюйгенса, принципу Ферма и т. д.

Для электромагнитных волн $F = c$ для всякой однородной и неоднородной среды. Значит, плоский фронт всегда остается плоским и идет со скоростью света в пустоте, какие бы тела ни стояли на пути волны. Эта теорема оправдывает применение методов геометрической оптики в расчетах, делаемых при сейсмических методах геологической разведки по приходу переднего фронта волн.

Л. И. Мандельштам. Закон дисперсии электромагнитных волн дает $F = c$. Можно ли положиться на этот вывод? Ведь сам закон дисперсии перестает быть верным при $k \to \infty$, ибо он выводится в предположении, что на протяжении одной длины волны имеется еще очень много диполей. Но изложенная теорема позволяет, по-моему, сказать больше. Она говорит, что мне не надо знать закона дисперсии, а надо знать только скорость очень короткой монохроматической волны. Действительно, как колеблется один осциллятор под действием синусоидальной силы одной частоты, я знаю. Я не знаю только, как мне ввести действие осцилляторов в показатель преломления в случае очень коротких волн. Но этого мне теперь и не нужно. При $\omega \to \infty$ осциллятор, наверно, не колеблется, значит, волна с $\omega = \infty$ проходит так, как если бы

электронов не было вовсе, т. е. без всякой дисперсии. Конечно, для этого вывода необходимо, чтобы дифференциальные уравнения (1) оставались в силе и для $\omega \to \infty$.

То, что фронт электромагнитной волны ни при каких условиях не может идти быстрее c, довольно ясно из обычных простых и правильных соображений насчет времени раскачивания электронов. Я сказал бы наоборот, что из этих простых соображений отнюдь не сразу видно, что фронт не может идти со скоростью, меньшей c, т. е. невозможно такое положение, что первичная волна, интерферируя со вторичными (все они идут со скоростью c), гасится вплоть до некоторой точки, движущейся медленнее, чем c.

Вопрос о характеристике был развит здесь для уравнений первого порядка. Очень часто мы имеем дело с уравнениями второго порядка. Конечно (с известными оговорками), их можно привести к уравнениям первого порядка, но можно формулировать изложенную теорему в несколько измененном виде — непосредственно для случая уравнений второго порядка. А именно: если на отрезке какой-либо кривой заданы φ_m и их производные, скажем по нормали к этому отрезку, то опять решение определяется однозначно при условии, что исходная кривая не является характеристикой. С этим связано деление уравнений второго порядка на эллиптические, гиперболические и параболические.

Следующее обстоятельство кажется здесь на первый взгляд странным. Мы говорим, что можно задать произвольно сами функции и их производные по нормали и тогда решение определится однозначно. Но ведь в ряде задач, в частности в теории потенциала, решение однозначно определяется заданием значений *самих функций*, а задания производных *не нужно*. Изучая дифракцию, мы даже ставили в минус теории Кирхгофа то, что она задает и φ и $\frac{\partial \varphi}{\partial n}$, в то время как достаточно задания φ, и тогда $\frac{\partial \varphi}{\partial n}$ уже *нельзя* задавать. В чем же здесь дело? На простом примере можно показать, что оба утверждения вполне совместимы. М. А. Леонтович показал, что можно построить требуемый ряд, но будет ли он всюду непрерывен и будет ли давать непрерывные производные? Это может быть не так, да мы этого и не требовали. В тех случаях, когда мы ищем решение по заданию только самих функций, мы требуем, кроме того, однозначности и непрерывности решения. Только этим обеспечивается его единственность.

Предположим, что мы ищем решение уравнения Лапласа в двух измерениях, причем зависящее только от r (осевая симметрия). Мы имеем

$$\varphi = a \ln r + b.$$

Если я требую, чтобы на круге $r = R$ имело место $\varphi = C$ и чтобы

φ всюду было непрерывно и однозначно, то единственное решение будет

$$a = 0,\ b = C,\ \text{т. е. } \varphi = C \text{ всюду.}$$

Если же мы хотим, чтобы не было особенностей только на круге $r = R$, то можно задать и φ_R и $\left(\frac{\partial\varphi}{\partial r}\right)_R$, определять a и b, и тогда при $r \to 0$ всегда будет особенность.

Еще один любопытный вопрос. Ведь у сигнала нули имеются, вообще говоря, не только на фронте, но и в других точках позади фронта. Но значит ли это, что они должны идти с той же скоростью, что и фронт, т. е. для электромагнитных волн со скоростью c? Легко видеть, что для любых нулей это не так. Волновое уравнение — второго порядка, и поэтому на фронте должна быть нулем и волновая функция и ее производные. *Только для таких* нулей скорость будет c. Следовательно, это не относится к нулям обычной синусоиды. Но возьмем монохроматическую волну в недиспергирующей среде. Мы имеем волновое уравнение, для которого постоянная всегда является решением. Добавив к синусоиде постоянную (равную амплитуде, рис. 7), мы через каждую волну получаем точку, где и функция и производная равны нулю. Это уже страшно: а что, если есть дисперсия? Тогда *все* нули идут со скоростью c, т. е. все синусоиды должны двигаться со скоростью c, а не $v(k)$.

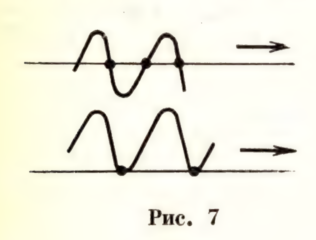


Рис. 7

Ошибка вот где: добавлять постоянную можно только тогда, когда в уравнения входят производные, но не сами функции. Если в уравнения входят сами функции, то постоянная уже не будет решением. Но в диспергирующей среде уравнения обязательно содержат и сами функции, т. е. $c_{jm} \neq 0$ в (1) (согласно (3), дисперсия есть только тогда, когда хотя бы один из коэффициентов c_{jm} отличен от нуля).

Еще одно замечание: с какой скоростью движется фронт ψ-волны в волновой механике? Если мы работаем с уравнением Шрёдингера, то связь между энергией и импульсом частицы такова:

$$E = \frac{p^2}{2m}.$$

Переход к дебройлевским волнам $\left(\omega = \frac{E}{h},\ k = \frac{p}{h} = \frac{mv}{h}\right)$ дает

$$v = \frac{\omega}{k} = \frac{hk}{2m},$$

и, следовательно, $F = \infty$, что противоречит принципу относительности. Конечно, это противоречие ничего не значит, так как уравнение Шрёдингера релятивистски не инвариантно. С уравнениями Дирака все получается правильно. Связь энергии и импульса здесь сложнее:

$$\frac{E^2}{c^2} = m^2c^2 + p^2,$$

откуда

$$c = \frac{\omega}{k} = c\sqrt{1 + \left(\frac{mc}{hk}\right)^2}.$$

Таким образом, здесь $F = c$.

СЕМИНАР ПО ОТДЕЛЬНЫМ ФИЗИЧЕСКИМ ПРОБЛЕМАМ

(1938—1939 гг.)

ВСТУПИТЕЛЬНАЯ ЛЕКЦИЯ К СЕМИНАРУ 1938/1939 г.

(19.XI 1933 г.) [1]

Я хотел бы коротко очертить тематику семинара и ее мотивировать. Нечего говорить, это все хорошо знают, насколько велики сдвиги, происшедшие в физическом мировоззрении за последние 25 лет. Эти сдвиги вылились в два учения: с одной стороны, специальную и общую теорию относительности, с другой стороны, квантовую теорию.

Первый вопрос, который здесь, естественно, возникает, состоит в том, каково же соотношение между новыми воззрениями и классической физикой. Конечно, здесь нет двух соперничающих физических систем. Положение не таково. Новые теории претендуют на то, что именно они правильно описывают явления. Однако это не значит, что классика потеряла свое значение. Дело в том, что новые теории — это не спекуляция, не выдумка досужих людей, а теории, которые выросли из потребностей опыта. С другой стороны, и классическая физика выросла из фактов, и если она долгое время нас удовлетворяла, то это не могло быть случайностью.

Вы, конечно, знаете, каково действительное положение вещей. Современная физика признает компетенцию классики, но, говоря

[1] [Семинар 1938/39 учебного года был посвящен отдельным физическим вопросам, выбор которых мотивирован в помещенной здесь вступительной лекции Л. И. Мандельштама. На семинаре были сделаны доклады: 1) О поле электрона, движущегося в среде со скоростью, превышающей фазовую скорость света в этой среде, и об эффекте Черенкова (И. Е. Тамм). В этом докладе были рассмотрены также соответствующие задачи об излучении быстро движущихся тел в акустическом случае и в случае волн на поверхности жидкости; 2) Об эффекте Допплера; 3) О чисто механическом обосновании закона эквивалентности массы и энергии и о выводе лоренцова преобразования на основе этого закона и принципа относительности (студент Г. Друкарев). К этому докладу Л. И. Мандельштам сделал несколько дополнительных замечаний; это выступление также помещено ниже. Обработка вводной лекции и указанного выступления выполнена С. М. Рытовым по собственным записками.]

схематически, ограничивает ее в двух направлениях: требованием скоростей, гораздо меньших c — скорости света в пустоте, и требованием достаточно больших масс, размеров и т. п. Классическая физика развивалась на макроскопическом опыте и переносила свои закономерности в микрокосмос. Современная физика считает, что этого как раз и нельзя делать.

Таким образом, классика выступает сегодня как предельный случай новых, более общих теорий. Но когда, где и как происходит переход?

В современных курсах все это излагается, но, естественно, при изложении стараются придать теории общий, логически законченный вид и тем самым указывают на предельный переход к классическим законам в очень общей и абстрактной форме. Поэтому, хотя предельные переходы известны и ясны, но не всегда достигается ощущение этих вещей, не всегда достигается и чувство необходимости новых воззрений. Между тем знать и овладеть — это разные вещи. Поэтому мне казалось желательным выбрать несколько по возможности фундаментальных, но вместе с тем и возможно более простых и цельных вопросов и на них показать, как и почему с необходимостью происходит переход к новым воззрениям, тут же подкрепив теоретическое рассмотрение соответствующими экспериментальными фактами. Это составит первый раздел семинара.

Какие же именно вопросы войдут в этот раздел? Я остановился на трех действительно фундаментальных (частью не только для теории, но и для эксперимента) вопросах.

Первая тема, правда, несколько нарушает изложенный план, ибо в ней речь будет идти о явлениях, теория которых во всех существенных чертах ясна и достаточно точна уже с классической точки зрения.

Когда господствовала лоренцова теория электромагнетизма, ставился вопрос о том, как ведет себя заряженная частичка, движущаяся с определенной скоростью (в частности, постоянной). Решение было найдено, и оказалось, что при $v \to c$ оно теряет смысл. Тогда не было видно, почему c играет какую-то особую роль, почему нельзя взять $v > c$. Последний вопрос был поставлен и решен Зоммерфельдом. Решение оказалось забавным: в то время как для $v < c$ поддержание постоянного значения v не требует приложения внешней силы, для $v > c$ нужна сила, т. е. имеется сопротивление и равномерное движение уже не может быть движением по инерции. Зоммерфельд нашел эту реакцию и правильно приписал ее наличию излучения. Таким образом, электрон, движущийся с постоянной скоростью $v > c$, излучает, но Зоммерфельд тут же добавляет, что он не знает, как происходит самый переход через $v = c$.

Теория относительности утверждает, что в природе нет скоростей, бо́льших c. Это одно из основных ее утверждений. Тем самым решенная Зоммерфельдом задача потеряла смысл, превратилась просто в математическую задачу. Однако этот вопрос в ином виде возродился теперь вновь. Несколько лет назад П. А. Черенков обнаружил очень слабое, но интересное явление — свечение чистых жидкостей под действием γ-лучей[1]. С. И. Вавилов[2] указал на то, что это не обычная флуоресценция (излучение не тушится тушителями, поляризовано преимущественно вдоль первичных γ-лучей, в отличие от обычной флуоресценции, либо неполяризованной, либо, если ее вызывает поляризованный свет, поляризованной преимущественно так, как и этот первичный свет). Дальнейшие опыты (с магнитным полем) показали, что свечение обусловлено не самими γ-лучами, а вырываемыми ими комптоновскими электронами. И. Е. Тамм и И. М. Франк дали очень изящную теорию этого явления[3].

Действительно, $v > c$ невозможно. Но фазовая скорость света в среде есть $c/n < c$, и, следовательно, в среде электрон *может* двигаться со скоростью $v > c/n$. Конечно, c/n, вообще говоря, неопределенная величина, ибо $n = n(\omega)$, т. е. существует дисперсия и только для $\omega \to \infty$ мы во всех телах имеем $n \to 1$. Это обстоятельство — наличие дисперсии — учтено в теории Тамма и Франка, которые показали, что для $v < c/n$ электрон ведет себя обычным образом, но если для какой-либо $\lambda v > c/n(\lambda)$, то он начинает излучать эту длину волны.

Таким образом, здесь опять выплыл наружу электрон, летящий быстрее света, но света в среде, т. е. без противоречия с теорией относительности. Все получается здесь очень элегантно и, быть может, показывает, что рано было забрасывать задачу Зоммерфельда и ждать, пока эксперимент натолкнет на это явление.

Недавно в Phys. Rev. появилось сообщение американских авторов, которые проверили и продолжили опыты Черенкова с лучшими средствами. Опыты сделаны хорошо, авторы весьма лояльны по отношению к Черенкову, но теории явления, ее гвоздя, они не поняли совершенно. Они говорят, что это то же самое, что и излучение рентгеновских лучей при торможении, не видя того, что здесь *равномерно* движущийся электрон.

Вторая тема тоже касается вещей, о которых говорилось немало странного — принципа Допплера (1842 г.). Уже в 1848 г. Физо указал, что эффект Допплера пригоден для измерения луче-

---

[1] [*П. А. Черенков.* ДАН СССР, 2, 451, 1934.]
[2] [*С. И. Вавилов.* ДАН СССР, 2, 457, 1934.]
[3] [*И. Е. Тамм, И. М. Франк.* ДАН СССР, 14, 107, 1937. См. также «Труды Физического института им. П. Н. Лебедева», т. II, вып. 4, 1944.]

вых скоростей звезд. Этот эффект играет большую роль и в наземных явлениях — в акустике, ультраакустике, для каналовых лучей (Штарк), в рассеянии света.

С точки зрения классической лоренцовой физики есть два различных случая: 1) наблюдатель покоится в эфире и движется источник и 2) источник покоится в эфире и движется наблюдатель. В обоих случаях имеет место смещение ω, но расчет и интерпретация различны. Принцип относительности считает оба эти случая за один: имеет место *относительное* движение источника и наблюдателя. Этот вопрос действительно прост — речь идет о кинематике: есть среда, есть источник, спрашивается, как должен быть произведен пересчет от одной системы отсчета к другой. Правда, именно с этой точки зрения для классики здесь имелись затруднения, но об этом после.

Вместе с тем этот вопрос и фундаментален. Здесь неважно ни то, что речь идет именно об электромагнитных волнах, неважна природа источника и т. д. Здесь легко понять особенности релятивистской кинематики, ибо речь идет именно о кинематике. Поэтому всякая теория пробует себя в первую очередь на принципе Допплера, и Эйнштейн уже в первой своей работе после общих рассуждений смотрит, как у него выходит принцип Допплера.

Эйнштейн получил результат, отличающийся от классического, хотя и во втором лишь порядке (впрочем, классически второй порядок не был рассчитан). Между тем источник, движущийся настолько быстро, что различие между классикой и теорией относительности уже измеримо современными средствами, имеется. Это — каналовые лучи. Появилась работа Айвса, опыты которого однозначно говорят о том, что в природе существует именно релятивистская кинематика. Между тем в этой работе имеются ссылки на предшествующие *теоретические* работы и нет ни слова о принципе относительности, из которого все так просто выходит. В чем дело — непонятно.

Третья тема касается еще более фундаментального вопроса теории относительности. Один из ее результатов, самый важный и плодотворный — эквивалентность массы и энергии. Именно ввиду общности этого положения его и обосновывают очень общим образом — из электромагнитных соображений. Это естественно, поскольку сама теория относительности выросла из электродинамики. При переходе же к механике закон эквивалентности, закон чрезвычайно общий, обрастает рядом дополнительных предположений. Между тем этот закон важен сегодня именно в механике, в ядерной физике, в актах соударения частиц, в ядерных реакциях. Таким образом, представляет интерес снизить до минимума предпосылки и постулаты, необходимые и достаточные

для *механического* обоснования закона эквивалентности, являющегося теперь одним из основных законов физики.

Попытки такого обоснования делались неоднократно, с самого возникновения теории относительности (Толмен, Вайль). Несколько лет тому назад (1932 г.) Ланжевен дал очень простое, ясное, и отчетливое изложение чисто механического обоснования. Совсем недавно (1937 г.) такой вывод дал Паули. Здесь все становится ясным — и роль кинематики, и тип законов сохранения.

Но сегодня хотелось бы посмотреть на этот вопрос и с другой стороны. Ядерная физика окончательно подтвердила кинематику теории относительности. Считая закон $E = mc^2$ правильным, теперь уже поправляют химические веса, т. е. это дает бо́льшую точность, чем в состоянии дать химические методы. Но если опыт достаточен для того, чтобы полагать кинетическую энергию T не в виде $\frac{mv^2}{2}$, в виде $m_0\left(\frac{1}{\sqrt{1-\beta^2}}-1\right)$, то можно, обратно, спросить себя: какой тип кинематики совместим с этим выражением кинетической энергии? Этот вопрос следует разобрать. Мне кажется, что единственной пригодной кинематикой является эйнштейновская.

Таковы три темы первого раздела семинара.

При выборе тематики второго раздела я руководствовался существенно иной точкой зрения, и это лежит в существе дела. Всякая физическая теория, поскольку она является теорией, имеет две стороны. Возьмем пример.

Если мы хотим рассматривать колебания закрепленной на концах струны, то теория приводит к следующему уравнению:

$$\frac{\partial^2\varphi}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2},$$

причем при $x = 0$ и $x = l$, $\varphi = 0$.

Теоретическая физика занимается решением и дискуссией таких уравнений (написанное уравнение получается для антенны, для лехера, для органной трубы и т. п.). Интересно и существенно, что

$$\varphi = A\sin(\omega t - kx)$$

будет частным решением при определенных *дискретных* значениях k. Однако пока все это ничего общего не имеет с физикой. Это математика, это математические обозначения, которые еще надо связать с физическими величинами. Для этого недостаточно просто сказать, что x и φ — координаты, а t — время. Надо указать, *как измерять* эти величины, т. е. нужны твердые тела

(масштабы), часы, рецепты измерений. Если и это дано, то только тогда мы будем иметь дело с физической теорией. В классике подобная постановка вопроса представляла мало интереса: было непосредственно ясно, что и как измерять. И специальная теория относительности в этом отношении целиком классична. Она также мало занимается этими вопросами. Единая теория поля идет здесь уже несколько иначе, а волновая механика идет совсем иным путем, чем шла классика.

Главное, что больше всего поражало при возникновении теории квантов,— это дискретность (энергии, орбиты и т. д.). Я думаю, что Шрёдингер искал такой математический аппарат, который давал бы эту дискретность. И он нашел его в краевых задачах, получив ряд дискретных величин. Таким образом, здесь путь обратный: не от реальных объектов и соответствующих им физических величин к соотношениям между последними, а обратно — от математического аппарата к физике.

Шрёдингер нашел уравнение

$$\nabla^2\psi + \frac{2m}{h^2}(E - V)\psi = 0.$$

Вопрос о том, что такое x, ψ и т. д., вопрос об этой второй стороне, которая только и делает теорию физической, встал немедленно. После ряда колебаний и различных мнений Борн стал на статистическую точку зрения: $|\psi|^2 dx$ есть *вероятность* того, что координата x лежит в интервале $(x, x + dx)$. Но что такое *координата* в квантовой механике? Я это плохо знаю, и вряд ли кто-нибудь это знает. Таким образом, в волновой механике центр тяжести сместился на эту вторую сторону (в классике — тривиальную): как трактовать величины, как связать их с реальными объектами.

И вот я хочу рассмотреть во второй части семинара ряд эпизодических вопросов о том, как и что *измеряют* в квантовой механике. Здесь есть два характерных отличия от теории относительности. Во-первых, в отличие от теории относительности, в волновой механике какое-либо общее положение излагается обычно на частном примере (микроскоп, щель и т. д.). Во-вторых, в квантовой теории вопросы измерений, в отличие от теории относительности, дискуссионны по существу, в них нет окончательной ясности. Может быть, мы сумеем разобраться до конца, может быть, нет, но здесь сам предмет подлежит дискуссии, и дело не в том, что *мы* того или иного не знаем, а в том, что вообще еще нет единообразной и установившейся точки зрения.

11 Л. И. Мандельштам

ОБ ОБОСНОВАНИИ РЕЛЯТИВИСТСКОЙ КИНЕМАТИКИ

(14.III 1939 г.)

На предыдущих занятиях были рассмотрены некоторые общие соображения о чисто механическом обосновании закона $E = mc^2$, т. е. обосновании, не апеллирующем к электромагнитным явлениям, и о выводе лоренцова преобразования из принципа относительности и закона $E = mc^2$, рассматриваемого как независимо установленный опытный закон. Можно подойти к вопросу о релятивистской кинематике и еще иначе. Возьмем лоренцово преобразование

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}, \qquad t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Это частный вид преобразования, соответствующий специальному выбору направлений координатных осей и их начального расположения, но он содержит основное ядро всей релятивистской кинематики. В преобразование входит один-единственный параметр v — относительная скорость систем отсчета K и K', направленная вдоль общей оси x этих систем. Если теперь от системы K' мы перейдем с помощью аналогичного преобразования к системе K'', движущейся относительно K' со скоростью v_1, и затем посмотрим, как связаны x'', t'' с x и t, то оказывается, что эта связь дается опять лоренцовым преобразованием с некоторым новым значением параметра. Говоря более определенно, оказывается, что совокупность лоренцовых преобразований образует *группу*. Понятие группы — одно из самых общих и плодотворных понятий математики, а указанное свойство совокупности лоренцовых преобразований — одно из чрезвычайно важных и интересных их свойств.

Группой называется совокупность любых элементов, для которых выполнен ряд следующих требований. Во-первых, дан рецепт составления из двух элементов третьего, причем последний обязательно должен также принадлежать к совокупности. Этот рецепт принято называть «умножением», а третий элемент — «произведением», хотя рецепт может не иметь ничего общего с обычным умножением. Для «произведения» должен выполняться сочетательный закон, т. е. $A(BC) = (AB)C$. Во-вторых, в числе элементов совокупности должна иметься «единица», умножение на которую любого другого элемента A дает в результате A. Наконец, для каждого элемента A совокупность должна содержать

обратный элемент A^{-1}, т. е. такой, который при умножении на A дает «единицу». Если все эти условия выполнены, то совокупность называется группой. Такие совокупности могут быть конечными и бесконечными, непрерывными и дискретными.

Все это чистая математика; но легко убедиться в том, что эти математические требования, как перчатка к руке, подходят к нашим физическим требованиям. Возьмем, однако, сначала какой-нибудь простой пример, скажем, совокупность чисел 0, 1, 2, ..., 9. Если «произведением» мы назовем арифметическую сумму двух элементов, то совокупность не будет группой: «произведение» $3 + 9 = 12$ не принадлежит совокупности. Но, назвав «произведением» арифметическую сумму без целого числа десятков, мы, как легко видеть, получаем группу. Нуль играет роль «единицы», элементы 1, 2, 3, 4, 5 являются обратными элементам 9, 8, 7, 6, 5. Но стоит нам только при указанном рецепте «умножения» откинуть нуль, и совокупность снова перестанет быть группой. Требование, чтобы совокупность была группой,— довольно жесткое требование, накладывающее на совокупность большие ограничения.

Так вот, совокупность лоренцовых преобразований $L(v)$, как сказано, образует группу. «Произведением» элементов является последовательное применение преобразований $L(v_1)\ L(v_2)$. Оно представляет собой опять лоренцово преобразование $L(v_3)$, причем

$$v_3 = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}\ .$$

Группа является непрерывной (v может принять любое значение от $-c$ до $+c$), единичное преобразование соответствует $v = 0$, а обратное преобразование — замене v на $-v$. Что это значит физически?

Физически мы *должны* требовать, чтобы наши преобразования составляли группу, ибо физически необходимо, чтобы непосредственный переход от системы отсчета K к системе отсчета K'' давал то же самое, что и переход $K \to K' \to K''$ при любой системе K'. В частности, переход $K \to K' \to K$ должен оставлять нас в системе K, т. е. быть эквивалентным преобразованию $L(0)$.

Поставим теперь в известном смысле обратный вопрос. Из требования однородности и изотропности пространства — отсутствия выделенных положений и направлений — мы устанавливаем, что переход от одной системы отсчета к другой должен осуществляться какими-то *линейными* преобразованиями. Мы требуем далее *выполнения принципа относительности* и, наконец, требуем, чтобы совокупность искомых преобразований составляла *группу*.

Что дают эти условия? Оказывается, что они почти однозначно определяют вид преобразования. Из них следует

$$x' = \frac{x - vt}{\sqrt{1 - \alpha v^2}}, \qquad t' = \frac{t - \alpha v x}{\sqrt{1 - \alpha v^2}}, \tag{1}$$

где α — произвольная, но *постоянная* величина, не зависящая от v.

Исторически развитие шло не так. В качестве исходных постулатов Эйнштейн взял принцип относительности и принцип независимости скорости света от движения источника. Очень часто теорию относительности упрекали за то, что она строится на недостаточном опытном материале, что опыт Майкельсона не является решающим основанием, так как его можно объяснить и иначе. Конечно, был ряд и других экспериментальных оснований (соображения де-Ситтера о двойных звездах и др.), но в целом экспериментальная база многими рассматривалась как недостаточная.

Посмотрим, однако, что означает теория относительности для механики. Ясно, что, поскольку преобразования другие, постольку и механические величины тоже другие: масса зависит от скорости, энергия и импульс выражаются иначе и т. д. Что здесь имелось в области опыта? Единственным подтверждением было поведение электронов в электрическом и магнитном полях, показавшее, что $m = m_0 \Big/ \sqrt{1 - \frac{v^2}{c^2}}$. Но ведь это тоже не механика, а электромагнитное явление. И вот в этом отношении — почему я и считал интересным остановиться на этих вопросах — сегодня дело радикально переменилось. Сегодня вопрос стоит так: либо вы отбросите все ядерные явления, будете их игнорировать, либо же вы примете во внимание весь этот огромный экспериментальный материал, прекрасно укладывающийся в теорию относительности. Теперь, и не привлекая электромагнитных явлений, для которых согласие с теорией относительности всегда было полным, нельзя говорить, что опытного подтверждения нет. Оно есть.

Требование группы привело нас к преобразованию (1). Если $\alpha = 0$, то масса не должна зависеть от скорости. Опыт же говорит, что это неверно. Ядерные явления требуют эквивалентности массы и энергии, требуют зависимости массы от скорости. Значит, $\alpha \neq 0$, и весь вопрос сводится к тому, чему равно α, т. е. к прямому измерению коэффициента в соотношении $m = \alpha E$. Так как этот коэффициент оказывается равным $1/c^2$, то независимость скорости света от движения источника получается теперь уже как следствие новых исходных положений.

ЛЕКЦИИ
ПО ОСНОВАМ КВАНТОВОЙ МЕХАНИКИ
(ТЕОРИЯ КОСВЕННЫХ ИЗМЕРЕНИЙ)
(1939 г.)

ПЕРВАЯ ЛЕКЦИЯ
(26.III 1939 г.)[1]

Остающееся у нас в этом году время я хотел бы посвятить разбору некоторых вопросов, относящихся к волновой механике. Позвольте предпослать этому разбору несколько совершенно общих замечаний.

Нас будет интересовать не математическая сторона дела, не решение конкретных задач. Мне кажется, что желательнее обсудить некоторые принципиальные вопросы, касающиеся основ волновой механики. Дело это благодарное, но нелегкое. Разумеется, я не могу взять на себя задачу систематического исчерпыва-

---

[1] [В основу публикуемого текста положены записи лекций, сделанные С. М. Рытовым. Кроме того, использованы: а) черновики Л. И. Мандельштама, содержащие заметки, сделанные им в процессе подготовки лекций (в ряде случаев эти заметки весьма подробны, однако это не относится к 4-й и 5-й лекциям); б) совершенно неправленые и недостаточно квалифицированные стенограммы трех первых лекций, сами по себе непригодные для печати, но позволившие установить, что записи С. М. Рытова достаточно подробны и содержат все наиболее существенное (по объему записи составляют около половины стенограмм).

Вслед за лекциями по теории косвенных измерений (весна 1939 г.) Л. И. Мандельштам предполагал прочитать в качестве их продолжения цикл лекций о связи математического аппарата квантовой механики со статистическим ее толкованием, о каузальности и т. п., положив в основу разбор работы Неймана (*F. v. Neumann.* Matematische Grundlagen der Quantenmechanik. «J. Springer», 1932). Обнаруженный среди рукописей Л. И. Мандельштама текст, который является, быть может, наброском первых лекций намечавшегося цикла, публикуется здесь в виде «Дополнения». Указанный черновик не закончен и обрывается на полуслове. Ввиду его тесной связи с лекциями по теории косвенных измерений (многое в этом наброске развивает содержание 1-й лекции) он помещен здесь, а не в III томе, где собраны остальные, найденные в бумагах Л. И. Мандельштама материалы.

Текст лекций и «Дополнения» обработаны Е. Л. Фейнбергом и просмотрены И. Е. Таммом и В. А. Фоком.]

ющего изложения. Это было бы чразвычайно трудно, и мое изложение по необходимости будет фрагментарным.

По-моему, концепция волновой механики многое отражает очень правильно, она показала себя весьма плодотворной. Но ряд основных понятий, которыми волновая механика оперирует, мне кажется, в основных изложениях трактуется весьма коротко, недостаточно ясно и не всегда убедительно. Больше того, у меня нет уверенности, что эти понятия уже вполне установлены, что они приобрели, например, ту четкость и ясность, которые, несомненно, присущи основным понятиям классической физики. И это, я думаю, не случайно.

То обстоятельство, что неясность и недоговоренность в основах волновой механики проистекают не из плохого изложения того или иного учебника, видно хотя бы потому, что эти основные вопросы дебатируются по сей день, и такими людьми, крупнее которых среди физиков мы не знаем. Полного согласия не было до последних лет между самыми большими авторитетами. Стоит вспомнить обмен мнений, происшедший несколько лет назад по вопросу о полноте волновой механики, в котором такие исследователи, как Эйнштейн, Бор, Шрёдингер и другие, высказывали прямо противоположные мнения [1].

Я вовсе не думаю, что нам удастся решить стоящие здесь вопросы до конца. На некоторые вопросы я затруднился бы дать сейчас (вполне) удовлетворительный ответ: тут просто еще нужна большая работа. В других, мне кажется, некоторые разъяснения дать можно, и вот о таких вопросах я хотел бы поговорить. Но и тут часто я не возьмусь делать утверждения с такой уверенностью, как в области классической физики, включая сюда и специальный принцип относительности, где основные понятия мне кажутся ясными. Но это не должно служить препятствием. Несмотря на это или, может быть, именно поэтому я считаю желательным на нашем семинаре такие вопросы поднять. Уже сама постановка их желательна; если же нам сообща удастся внести сюда несколько бóльшую ясность, то тем более нам стоит ими заняться.

И последнее замечание: основное знакомство с квантовой механикой мы, конечно, должны у слушателей предположить. Я думаю, что у большинства из вас оно есть.

Мне придется начать несколько издалека.

Какова структура всякой физической теории, всякого физического построения вообще? Немного схематично (как всегда) можно сказать, что всякая физическая теория состоит из двух дополняющих друг друга частей. Я начну с того, чтó можно счи-

---

[1] [См. лекцию 4.]

тать второй частью. Это *уравнения* теории — уравнения Максвелла, уравнения Ньютона, уравнение Шрёдингера и т. д. Уравнения — это просто математический аппарат. В эти уравнения входят некоторые символы: x, y, z и t, векторы **E** и **H** и т. д. На этом вторая часть заканчивается. Здесь еще нет никакой физической теории. Это математика, а не естественная наука. Первую же часть физической теории составляет связь этих символов (величин) с физическими объектами, связь, осуществляемая по *конкретным* рецептам (конкретные вещи в качестве эталонов и конкретные измерительные процессы — определение координаты, времени и т. д. при помощи твердых масштабов, часов и т. д.).

Первая часть учит, как рациональным образом отнести к объектам природы определенные величины — большей частью в виде чисел. Вторая часть устанавливает математические соотношения между этими величинами. Тем самым, ввиду связи этих величин с реальными объектами, формулируются соотношения и между этими последними, что и является конечной целью теории.

Без первой части теория иллюзорна, пуста. Без второй вообще нет теории. Только совокупность двух указанных сторон дает физическую теорию.

Пусть, например, я имею уравнение

$$\frac{d^2x}{dt^2} = g. \qquad (1)$$

Пока это — математика, и физику нечего здесь делать. Но мы говорим, что оно описывает поведение тела в поле тяготения, и утверждаем, что t — математический символ времени, т. е. мы утверждаем, что для нахождения значений t надо взять такие-то и такие-то часы и так-то ими пользоваться. Аналогично мы утверждаем, что x надо находить с помощью твердого масштаба и смотреть, с какой чертой на этом масштабе будет совпадать положение тела и т. д. Тогда математически найденный интеграл уравнения (1)

$$f(x, t, C_1, C_2) = 0 \qquad (2)$$

будет физическим законом, все пары измеренных x_i и t_i связаны этим законом. Если я узнаю, что $t = t_1$, и если C_1 и C_2 заданы начальными условиями, то я уже знаю, что $x = x_1$.

Таким образом, мы имеем вначале переход от объектов к числам при помощи рецептов, потом следует математика и затем уже, чтобы выразить полученный результат как физический факт, — обратный переход.

Конечно, указанные две стороны физической теории не независимы друг от друга. В частности, при введении определений необходимо должно отсутствовать противоречие между ними.

Так, например, в законы Ньютона, примененные, скажем, к движению планет, входят расстояния между тремя точками; с другой стороны, имеются рецепты, указывающие, как следует измерять эти расстояния. Но сами ньютоновские уравнения построены так, что в них предполагается, что сумма длин двух сторон треугольника больше третьей стороны: $r_1 + r_2 \geqslant r_3$. Допустим теперь, что я даю такой рецепт измерения r, при котором соотношение $r_1 + r_2 \geqslant r_3$ не соблюдается. Это значит тогда, что между математическими соотношениями теории и рецептами измерения существует противоречие, что указанный рецепт измерения непригоден.

Я приведу еще один пример из чистой классики, который гораздо больше, чем первый, важен нам для дальнейшего.

В оптике оперируют различными величинами, представляющими собой функции времени: $E = f(t)$. Кроме того, эта колебательная теория — волновая оптика — содержит частоту ν. Частота, по определению, дает периодичность в случае синусоидального изменения

$$E = \sin \nu t.$$

Для соотнесения величин E и ν конкретным физическим объектам мы даем два рецепта: 1) пусть луч света падает на пластинку, быстро движущуюся в своей плоскости перпендикулярно к лучу. Почернение пластинки, по определению, и есть измеренная в определенном масштабе величина E, выраженная как функция времени $f(t)$; 2) пусть луч света проходит призму, а затем линзу. Точка, получающаяся в фокальной плоскости линзы, по определению (посредством определенного пересчета из значений угла), дает частоту ν. И вот если функция $f(t)$ не есть синус, то здесь уже, по *определению*, нет частоты, или, лучше сказать, нет одной частоты. В соответствии с этим мы, поставив призму и линзу, получим теперь распределение точек в фокальной плоскости, описываемое некоторой функцией $g(\nu)$, которая и будет спектром функции $f(t)$.

Пусть $E = f(t)$ — острая функция, т. е. на движущейся пластинке получается короткое почернение. Если бы для того же пучка призма дала острое $g(\nu)$, то я сказал бы, что налицо имеется противоречие: либо неверно волновое представление о свете, либо же не годятся мои рецепты. Таким образом, связь характера функции $f(t)$ и характера ее спектра $g(\nu)$ заложена уже в определениях E и ν, т. е. на очень ранней стадии, еще до развернутой теории. Это справедливо и в максвелловской теории, и в акустике, и в теории упругости.

И вот этим условием, что некоторых противоречий заранее не должно быть, я связываю все построение теории очень сильно.

В рассмотренном примере, по определению величин $E = f(t)$ и ν, острое E необходимо должно давать широкий спектр.

Классическая физика большей частью шла так, что установление связи математических величин с реальными вещами предшествовало уравнениям, т. е. установлению законов, причем нахождение уравнений составляло главную задачу, ибо содержание величин заранее представлялось ясным независимо от законов. Мы просто свыклись с ними (возьмите, например, длину и т. д.) и для них искали уравнения. Нужно, однако, сказать, что и эта первая часть совсем не так проста, как часто кажется. Очень многие недоразумения в понимании основ даже специального принципа относительности основаны на том, что эта сторона физической теории недостаточно ясно усвоена.

Современная теоретическая физика, не скажу — сознательно, но исторически так оно было, пошла по иному пути, чем классика. Это получилось само собой. Теперь прежде всего стараются угадать математический аппарат, оперирующий с величинами, о которых или о части которых заранее вообще не ясно, что они означают.

Дело в том, что математический аппарат часто устанавливает очень характерные соотношения между параметрами независимо от того, каким объектам эти параметры соответствуют. Так, например, определенные дифференциальные уравнения устанавливают периодическую зависимость одних параметров от других. И вот, подмечая в физических явлениях, часто качественно, характерные особенности, ищут математический аппарат (т. е. строят вторую часть теории), который отражал бы эти характерные особенности, причем сначала не особенно заботятся о той связи, которая существует между всеми величинами, входящими в аппарат, и действительными объектами, и только потом стараются эту связь установить.

Так, несомненно, поступал Эйнштейн, особенно при создании общего принципа относительности. Это особенно ясно видно и на примере того, как создавалось уравнение Шрёдингера.

Вы знаете, что первый толчок всему современному развитию той теории, которую мы теперь называем волновой механикой, был дан Планком. Основным положением Планка, так резко идущим вразрез с принципами классики, было положение о дискретности возможных уровней энергии осциллятора. Блестящее развитие планковских идей привело Бора к его представлению об атоме, опять наиболее замечательной чертой которого было утверждение о дискретности энергетических состояний атома.

Дальнейший существенный шаг Бора вперед состоял в установлении связи между частотой (испускаемого света) и энергией, или, вернее, разностью энергий двух состояний.

Вот это наличие дискретных энергий, или частот, так блестяще объясняющее спектральные серии, опыты Герца — Франка и т. д., — именно это считалось тогда (оно замечательно и сейчас), так сказать, гвоздем теории.

И мне кажется, что не будет ошибкой сказать, что как де Бройль, так и Шрёдингер искали в первую очередь математический аппарат, т. е. искали такое построение математической стороны теории, которому было бы присуще выделение дискретных значений, мало заботясь на этой стадии о первой части, т. е. мало заботясь о том, какое физическое значение нужно будет приписать тем величинам, которые будут входить в этот математический аппарат, за исключением самих дискретных величин, которые, очевидно, надо было толковать как E [энергию] или ν [частоту] — ведь для этого все делалось.

И в поисках этого аппарата, несомненно, решающую роль сыграло то обстоятельство, что в классической физике такой аппарат существовал. Это так называемые задачи о собственных значениях, т. е. по существу волновые уравнения с граничными условиями. Ими решались задачи о собственных (дискретных) колебаниях струн и т. д. Там конечно, физический смысл всех математических символов ясен. Там, этот аппарат появляется при приложении общих принципов механики или электродинамики к частным вопросам.

Для краевых задач типично выделение целых чисел, обеспечивающих существование решения. То, что сделал Шрёдингер, что он угадал, замечательно вот в каком отношении. В струне дискретные числа выделены не столько уравнением, сколько граничными условиями. А где взять эти граничные условия для атома, за что там зацепиться? И вот Шрёдингер обратил внимание (математики знали это давно), что существуют такие дифференциальные уравнения, которым для появления дискретности достаточно *естественных* условий, таких, например, как интегрируемость квадрата модуля решения и конечность его в особых точках уравнения. Именно таким путем Шрёдингер пришел к своему уравнению, и он сам говорил, что не ожидал такого поразительного эффекта от своей работы.

Шрёдингер сразу же предположил, что добытые на этом пути дискретные числа надо связывать с энергией рассматриваемой системы.

Возьмем конкретный случай, например линейный осциллятор (а не атом водорода, как взял первым примером Шрёдингер). Мы имеем здесь уравнение

$$\nabla^2\psi + \frac{8\pi^2 m}{h^2}(E - V)\psi = 0, \tag{3}$$

где $V = \alpha x^2$. Требование конечности интеграла $\int\limits_{-\infty}^{+\infty} |\psi|^2 dx$ тотчас же дает нам, что уравнение имеет регулярные решения не для всех значений параметра E, а только для дискретного множества.

Чтобы уравнение Шрёдингера не было пустым, т. е. чтобы оно имело физический смысл, нужно связать входящие в него математические величины с физическими объектами. Естественно при этом, что мы постулируем: E есть энергия осциллятора — ведь для этого уравнение Шрёдингера и было создано.

Но в уравнение (3), кроме E, входят и другие величины, и в первую очередь x и функция ψ. И положение было таково, что, установив свое уравнение, Шрёдингер не знал, как связать *эти* величины с природой. Однако, еще не зная всего этого, зная лишь смысл E, он получил очень многое: серию Бальмера, штарк-эффект, нулевую энергию и т. д. Создалось странное положение, когда дебатировались остававшиеся неясными самые основы теории и между тем получались хорошие прикладные результаты. Конечно, можно было бы сказать, что этих результатов достаточно и вообще больше ничего не нужно. Однако вряд ли можно считать теорию достаточно полной, если не установлена связь между объектами и всеми величинами, входящими в теорию. Тогда же, с самого начала, развернулись поиски смысла функции ψ. Как видите, второй путь построения физической теории много труднее, но я думаю, что это вполне правомерный путь.

Часто делают упрек волновой механике, говоря, что это чистая математика. Это не совсем справедливо. Путь волновой механики действительно был таким, что сначала шла математика. Но я не вижу, чем этот путь хуже классического, *если в дальнейшем связь с природой найдена*. Во всяком случае, *до или после установления уравнений установление такой связи необходимо*, без этой связи само уравнение не может претендовать на название физической теории. И здесь действительно лежит корень того, что, я считаю, недостаточно ясно излагается в учебниках и в литературе и что повело к очень многим неясностям.

Начнем со следующего.

В настоящее время установили (или думают, что установили), что $|\psi|^2 dx$ — вероятность нахождения частицы в интервале $(x, x + dx)$. А что такое x? x — координата. Так вот, и в первом утверждении, и во втором есть некоторый самообман.

Можно, по-видимому, сказать следующее. Квантовая механика совершенно правильно рвет с предрассудком, что законы макромира действительны и в микромире. Но последовательно она про-

водит эту точку зрения (во всяком случае, в явном виде, в учебниках) только для второй части теории (математической) и не учитывает в достаточной мере, что и рецепты перехода тоже должны быть даны иными, чем в классике. Наша задача и будет состоять в том, чтобы выяснить, что имеют в виду в квантовой механике, когда говорят об измерении той или иной величины.

Если я в классике говорю, что x есть координата данной материальной точки, то за этим определением стоит вполне ясный, конкретный рецепт: если я установлю соответствующим образом твердый масштаб с нанесенными по определенному рецепту делениями, то x — это число на том делении, с которым в данный момент совпадает моя точка. Понятие «совпадает» для макромира считается известным. И поскольку это так, мы действительно установили рецепт для перехода от символа x к реальным объектам (материальная точка, твердое тело — масштаб и т. д.). Такого рода рецепты мы называем измерениями.

Поскольку, однако, речь идет о молекулярных вопросах, такие рецепты не выполнимы принципиально, а не только практически (*элементарность* понятия «совпадение» стоит под вопросом из-за воздействия прибора). Поэтому, назвав x координатой, я не установил связи x с природой, я только сделал вид, что установил эту связь, сославшись на макромир. С такими «определениями» теория еще висит в воздухе. Правильнее было бы даже и называть x не координатой, а, например, квазикоординатой.

Далее, говорят, что $|\psi|^2$ — вероятность. Допустим, что мы знаем, как измерить x (в действительности, на нашей стадии рассмотрения,— не знаем). Тогда создается впечатление, что смысл $|\psi|^2$ уже вполне ясен. Однако здесь имеется недоговоренность, и, прежде чем перейти к рассмотрению всех этих вопросов, нужно остановиться еще на одной стороне дела.

Волновая механика — статистическая теория. Но говорить о статистике и о вероятности можно, только имея определенную совокупность элементов, к которой эта статистика относится. В волновой механике такой совокупностью является совокупность повторных опытов (каждый индивидуальный опыт есть ее элемент), причем повторение должно происходить при одних и тех же условиях. Именно к такой совокупности относятся вероятностные высказывания, о которых говорилось выше. Необходимо подчеркнуть, что и всякая классическая теория по существу также является статистической, и в этом пункте нет принципиального отличия волновой механики. Если бы теория давала высказывания только для одного опыта, она не имела бы ценности. Но классика, например классическая механика, утверждала, что возможны сколь угодно узкие распределения для всех исследуемых

величин. Таким образом, разница — в *типе* статистики, а не в самой статистичности.

В обоих случаях, и в классической теории, и в волновой механике, мы имеем дело с большой совокупностью элементов и некоторым признаком. Назовем эту совокупность, над которой проделывается статистическая обработка, коллективом. Коллектив должен быть как-то выделен, иначе теряет смысл постановка любого вопроса о нем. Так вот, говорят, что $|\psi|^2$ — вероятность. Но в каком коллективе? Если это не указать, то возможны всякие неясности и парадоксы. «Средний рост человека такой-то»,— без указания коллектива это утверждение вообще не имеет смысла. Рассказывают, что однажды захотели определить среднюю величину семьи по результатам ответа на вопрос, сколько детей у ваших родителей. Очевидно, это не может дать средней общей величины, так как семья без детей автоматически не учитывается. Замечу, что полемика Эйнштейна и Бора вызвана именно грубой ошибкой в вопросе о коллективе [1]. Я не встречал достаточно четкого определения коллектива, в котором берется ψ-функция. Уточнить этот коллектив — первое, что нам нужно сделать.

Разумеется, и в классике мы сталкиваемся с тем же вопросом. Мы можем говорить о максвелловском распределении скоростей только при постоянной температуре. Если температура изменяется, то распределение будет совсем другое. То же самое имеет место и в тех классических задачах, в которых не говорят о коллективе. Я уже сказал, что любой опыт классики есть тоже статистический опыт, ибо только он имеет смысл. Пусть, например, мы наблюдаем движение маятника и ставим задачу уравнением $m\ddot{x} + \alpha x = 0$. При этом мы фиксируем α, т. е. предполагаем, например, что температура и ускорение силы тяжести не меняются. Таким образом, при всяком теоретическом рассмотрении *условия опыта* надо определить, и это определение всегда может быть сведено к фиксированию некоторых параметров.

Мы подошли к тому, что я считаю наиболее существенным и важным. А именно, волновая механика утверждает, что для *определения микромеханического коллектива, к которому и относится ψ-функция, достаточно указать (фиксировать) макроскопические параметры.*

Если, например, в эвакуированной трубке из накаленной нити, к которой приложено известное напряжение, летят электроны, то их поведение подчиняется волновой механике. Элементом совокупности, о которой здесь идет речь, является поведение отдельного электрона, например его попадание в ту или иную точку анода или экрана. К этим элементам относится определенная

---

[1] [Подробнее см. лекцию 4.]

функция ψ, т. е. этой функцией и задается статистика вылетающих электронов. Но для того, чтобы электронный коллектив был выделен, должны быть фиксированы напряжение, температура нити, ее конфигурация и т. д., т. е. определенные *макроскопические* параметры. Если вы нарушите один из параметров, то вы получите другое значение ψ или, может быть, вообще никакого ψ не получите.

Таким образом, предпосылкой применения волновой механики является экспериментальная установка, *контролируемая классическими измерениями.*

Теперь, после того, как мы определили свою микростатистику макроскопическими условиями, мы должны обратиться к вопросу о том, что же мы называем в квантовой механике координатой, а также импульсом, скоростью и т. д. Здесь волновая механика идет не обычным путем, но путем, чрезвычайно для нее характерным.

Из математических величин, входящих в уравнение Шрёдингера, она строит другие величины и утверждает, что именно для этих последних должны быть даны рецепты, устанавливающие связь с реальными вещами. Путь построения этих новых величин состоит в том, что каждой физической величине сопрягается соответствующий ей эрмитов оператор; операторы имеют свои собственные значения, и вот для этих-то собственных значений и должны быть даны измерительные рецепты.

Собственным значениям отвечают собственные функции оператора. Утверждается: если у вас есть некоторая функция ψ, которая не является собственной функцией оператора, то ее следует разложить по собственным функциям, и тогда, как утверждает волновая механика, квадраты модулей коэффициентов такого разложения дают вероятности того, что рассматриваемая величина даст при измерении соответствующие собственные значения.

Все это — общеизвестные математические правила. С их помощью образуются новые величины, которые и надлежит связать потом с объектами природы. Вот этого-то и не делают в обычных изложениях, т. е. не дают рецептов связи — не осуществляют второй части всей программы. Отсюда сразу же возникают все неприятности и неясности.

Для импульса p мы имеем оператор $\frac{\hbar}{i}\frac{\partial}{\partial x}$. Следовательно, для определения собственных значений p' оператора p мы имеем уравнение

$$\frac{\hbar}{i}\frac{\partial\varphi}{\partial x} = p'\varphi. \tag{4}$$

Это уравнение решается для любого p', так что собственными функциями будут функции

$$\varphi = e^{\frac{i}{\hbar} p'x}. \tag{5}$$

По только что высказанному постулату волновой механики это означает, что измерение может привести к любому значению p'. Вообще говоря, если волновая функция не совпадает с какой-либо из собственных функций, $\psi \neq \varphi$, то определенного импульса нет. Лишь в том случае, когда ψ — собственная функция p, система обладает определенным импульсом, а именно p'. Однако всякая функция ψ может быть разложена в интеграл Фурье, а тогда квадрат модуля коэффициента Фурье дает плотность вероятности каждого значения p' [1]. Заметим, что таким образом между p и x устанавливается точно такая же связь, как между ν и t в классике. Там мы уже видели, как связаны функция $f(t)$ и ее спектр $g(\nu)$, и здесь мы получаем то же самое: если ψ такова, что $|\psi(x)|^2$ дает очень острое распределение, скажем, около значения x', то уже невозможно, чтобы значение p' было каким-то определенным.

Значит, если мы так постулируем вид операторов, то еще до всяких уравнений Шрёдингера в самом определении величин, с которыми оперирует теория, исключается возможность одновременных точных значений x и p. Положение совершенно такое же, как в классике: «частота в данный момент» — это абсурд, независимо от того, справедливы ли максвелловы уравнения или какие-нибудь другие. Дело в определении самих величин. Обычно говорят, что соотношение неопределенностей возникает из-за взаимодействия измерителя и измеряемого объекта. Вы видите, что это соотношение возникает с самого начала, еще до вопроса об измерении. Однако взаимодействие измерителя и измеряемого объекта играет решающую роль в вопросе об адекватности величин, с которыми оперирует теория, реальным объектам. К этой стороне дела мы и перейдем, но в заключение — еще одно замечание по поводу предшествующего.

Почему мы называем p импульсом? Здесь ведь опять самообман: берется прежнее слово, и это создает видимость какого-то содержания, т. е. невольно, как и в случае координаты, подразуме-

---

[1] На первый взгляд может показаться, что x является исключением из описанной процедуры. Но это не так, x тоже является оператором. Им служит умножение на x. Таким образом, уравнение для определения собственных функций здесь имеем вид $x\psi(x) = x'\psi(x)$. Если ввести функцию Дирака $\delta(x - x')$, то она и будет собственной функцией x при собственном значении x'. Для любой $\psi(x)$ имеем тогда разложение по собственным функциям, в котором коэффициенты разложения из-за особых свойств собственных функций будут опять $\psi(x)$: $\psi(x) = \int \psi(x')\,\delta(x - x')\,dx'$, и по общему правилу получаем, что $|\psi(x')|^2$ есть вероятность того, что $x = x'$.

ваются прежние измерительные рецепты. Но пока у нас нет этих рецептов, и потому лучше всего было бы не пользоваться старыми терминами.

Соотношение неопределенностей нас потому и смущает, что мы называем x и p координатой и импульсом и думаем, что речь идет о соответствующих классических величинах. Называйте x и p квазикоординатой и квазиимпульсом. Тогда имеющееся между ними соотношение будет так же мало смущать, как соотношение между ν и t.

Итак, нам нужно теперь установить связь между этими математическими символами, входящими в уравнение Шрёдингера, и объектами природы. Для физика установить такую связь — это означает измерить, дать те конкретные рецепты, согласно которым из реальных вещей извлекаются численные значения теоретических величин. Эти рецепты могут быть сложными, состоящими из многих звеньев. В конечном счете это сводится к тому, чтобы дать эталон (например, масштаб) и способ его употребления (например, последовательное прикладывание). И то и другое нам понятно и близко лишь в макромире. Даже мысленно мы можем проделывать измерительные манипуляции только с макротелами. Я думаю поэтому, что *последнее звено необходимых нам измерительных рецептов обязательно макроскопическое.* Это мне кажется основным положением теории. Если бы дело обстояло не так, то я просто не вижу, как связать математические символы волновой механики с реальными объектами, а ведь без такой связи вообще нет физической теории. Поэтому, по крайней мере в настоящее время, мне кажется необходимым принять это положение.

Может случиться, что уже в первом звене мы имеем сведéние микровеличин на макрообъекты. Пусть, например, от нагретого катода летят электроны. Допустим, что состояние этих электронов описывается некоторой функцией ψ. Если электроны падают на пластинку и оставляют на ней пятна, то координату пятна я могу, по определению, считать координатой электрона. Статистика пятен должна описываться тогда функцией $\psi(x)$. И мое утверждение, что $|\psi(x)|^2$ есть вероятность нахождения электрона в точке x, приобретает теперь вполне конкретный смысл, так как величина x мною определена при помощи классических объектов. Такого рода случаи, когда измерение происходит в первом же звене, можно назвать *прямыми измерениями.* Но таких прямых опытов очень мало, и мне кажется, что это принципиальный и глубокий момент. Я думаю, что прямые измерения вообще возможны лишь для *свободных* или почти свободных частиц в слабых полях.

Но мы хотим толковать символы x и ψ также и в случае водородного атома. В нем не поставить фотопластинку, да и не в том

дело, что это практически неосуществимо. Я думаю, что в нем она могла бы и не почернеть. Поэтому я здесь лишен возможности дать определение координаты частицы на основании классических измерений. В таких случаях мы можем производить *косвенные* измерения, применяя некоторые посторонние частицы. Поэтому в рецепт войдут промежуточные звенья, которые, разумеется, могут быть не макроскопическими. Я должен, как будет видно ниже, идти по этому пути, пока не дойду до такого этапа, где стык с классикой несомненен. Но в конечном этапе вспомогательные частицы должны оказаться свободными.

Конечное звено всякого измерения есть классическое измерение потому, что оно должно сводиться к констатации совпадения «материальных точек» (деления на масштабе, стрелки и т. п.). Неудовлетворительность такого положения вещей заключается в том, что общего рецепта для указанного перехода я дать не могу. Отсюда видно, насколько существенно изучить косвенные измерения.

Итак, поставив вопрос о том, как же волновая механика устанавливает связь математических символов с реальными объектами, мы пришли к задаче — дать теорию косвенных измерений. Это необходимо, так как мы уверены, что прямые измерения являются исключением и что их исключительность не случайна, а имеет глубокий принципиальный характер.

ВТОРАЯ ЛЕКЦИЯ

(8.IV 1939 г.)

Напомню некоторые положения, которые мы установили на прошлой лекции.

Последнее звено измерения в волновой механике обязательно макроскопично. *Прямыми* измерениями мы называем такие измерения, в которых первый же шаг макроскопичен.

Принцип косвенного измерения состоит в том, что данную систему, в которой мы хотим измерить величину A, мы заставляем взаимодействовать с другой микросистемой, для которой уже возможно прямое измерение, и потом теоретически заключаем о значении A. Здесь взят случай прямого измерения во втором звене, но, вообще говоря, может быть сколько угодно промежуточных звеньев. Лишь последнее должно быть макроскопическим. Для

простоты я предположу, что уже во втором звене возможно сведе́ние на классическое измерение.

Здесь, по существу, входит рассмотрение двух взаимодействующих систем. Поэтому, прежде чем мы приступим к основному вопросу о косвенных измерениях, мы должны сделать довольно длинное отступление и рассмотреть сначала некоторые общие положения волновой механики, высказываемые ею для случая двух систем или одной системы, состоящей из двух частей.

Пусть имеются две системы, характеризуемые для простоты каждая одной «координатой», например две частицы, движущиеся по одной прямой. Пусть их «координаты» обозначены буквами x и y, $\Psi(x, y)$ — их волновая функция. В квантовой механике постулируется, что

$$dw = \Psi\overline{\Psi}dx'dy' \tag{6}$$

есть вероятность того, что одновременно x и y заключены соответственно в пределах $(x', x' + dx')$ и $(y', y' + dy')$.

Упростим терминологию. Если спектры собственных значений дискретны, то $\Psi\overline{\Psi}$ дает вероятность равенств: $x = x'$, $y = y'$. Если спектры непрерывны, то $\Psi\overline{\Psi}$ есть плотность вероятности неравенств $x' \leqslant x \leqslant x' + dx'$, $y' \leqslant y \leqslant y' + dy'$. Ради краткости мы и в этом случае будем говорить просто о вероятности $x = x'$ и $y = y'$, понимая под этим указанные неравенства.

Как найти вероятность того, что некоторая величина R, относящаяся к первой системе (характеризуемой координатой x), имеет одно из возможных своих значений λ_i? Постулируется, что для этого следует разложить $\Psi(x, y)$ по собственным функциям $\psi_i(x)$ оператора величины R

$$\Psi = \sum_i c_i\psi_i, \tag{7}$$

и тогда $c_i\bar{c}_i$ будет искомая вероятность.

Какова вероятность того, что величина R, относящаяся к первой из рассматриваемых систем (к системе I), имеет значение λ_i, а величина S, относящаяся к системе II, одновременно имеет значение μ_k? Искомая вероятность определится из разложения

$$\Psi(x, y) = \sum_{i,k} a_{ik}\psi_i(x)\varphi_k(y), \tag{8}$$

как $a_{ik}\bar{a}_{ik}$, причем ψ_i и φ_k — собственные функции операторов R и S, соответствующие собственным значениям λ_i и μ_k.

Пусть имеются две *независимые* системы, скажем, катодный пучок и пучок α-частиц. Для каждой системы можно написать свое уравнение Шрёдингера со своей волновой функцией [$\psi(x)$

и $\varphi(y)$]. Но я могу, по крайней мере мысленно, рассматривать эти две системы как одну с волновой функцией $\Psi(x, y)$. Какова эта общая волновая функция, если известно, что для одной из систем она есть $\psi(x)$, а для другой $\varphi(y)$? Как известно, если я хочу выразить с помощью Ψ то же самое, что выражают ψ и φ, то я должен положить

$$\Psi(x, y) = \psi(x)\varphi(y). \tag{9}$$

Разумеется, это надо показать, т. е. надо показать, что, не обращая внимания на α-частицы (на систему *II*), я буду получать для электронов (для системы *I*) с помощью Ψ те же физические результаты, что и с помощью отдельно взятой функции $\psi(x)$.

Пусть и для ψ, и для φ возможны прямые измерения координаты (например, по пятнам на пластинке, получающимся для электронов в одном месте, а для α-частиц — в другом). Тогда правильность высказанного утверждения очевидна: статистика для координаты x системы *I*, рассматриваемой независимо, дается функцией $|\psi(x)|^2$. Она получается совершенно такой же и из формулы (9). Действительно, вероятность значения x, независимо от того, чему равно y, получим, проинтегрировав $|\Psi(x, y)|^2$ по всем возможным значениям y. В силу нормированности функции φ мы опять получаем $|\psi(x)|^2$. Это для координаты.

Однако тождественность результатов должна иметь место для любой физической величины R. Разлагая $\psi(x)$ по собственным функциям ψ_i оператора R, $\varphi(y)$ — по собственным функциям φ_k, вообще говоря, иного оператора S, мы имеем

$$\left.\begin{aligned} \psi(x) &= \sum_i a_i\psi_i(x), \\ \varphi(y) &= \sum_k b_k\varphi_k(y), \\ \Psi(x, y) &= \sum_{i,k} a_i b_k \psi_i\varphi_k. \end{aligned}\right\} \tag{10}$$

Вследствие нормированности функций ψ_i и φ_k, очевидно, $\sum_i a_i\bar{a}_i = \sum_k b_k\bar{b}_k = 1$. По определению, $a_i b_k \bar{a}_i \bar{b}_k$ есть вероятность того, что система *I* находится в состоянии i, а система *II* — в состоянии k. Но нас интересует состояние системы *I* независимо от того, каково состояние системы *II*. По теореме сложения вероятностей получаем

$$w = a_i\bar{a}_i \sum_k b_k\bar{b}_k = a_i\bar{a}_i. \tag{11}$$

Это справедливо для любого оператора R. Если, с другой стороны, мы рассматриваем только изолированную систему I, то с помощью ее функции $\psi(x)$ получаем тот же результат сразу.

Но этим совпадением вопрос еще не исчерпывается. Это совпадение имеет место в начальный момент, когда системы I и II еще разобщены. Таким образом, для начальных Ψ, ψ и φ мы по доказанному имеем

$$\Psi = \psi\varphi \quad \text{при } t = t_0. \tag{12}$$

Но затем мы обращаемся с этими начальными функциями к уравнению Шрёдингера, с одной стороны, для ψ и φ, взятых порознь, с другой — к совокупному уравнению для Ψ. Каково будет соотношение между ψ, φ и Ψ в дальнейшем?

Если между системами I и II нет взаимодействия, энергия полной системы равна сумме энергий систем I и II, $E = E_1 + E_2$, то переменные разделяются, т. е. соотношение (12) остается в силе для *любого момента времени:*

$$\Psi(x, y, t) = \psi(x, t)\,\varphi(y, t).$$

В этом случае оказывается, таким образом, действительно безразлично, рассматривать ли две невзаимодействующие системы совместно или порознь.

Но возможны и другие, более интересные случаи.

В самом деле, уравнение

$$\frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2} - \frac{2m}{\hbar^2}\left[V_1(x) + V_2(y)\right]\Psi = -\frac{2mi}{\hbar}\frac{\partial\Psi}{\partial t} \tag{13}$$

имеет решения и не сводящиеся к виду (12), т. е. к одному произведению двух функций $\psi(x)$ и $\varphi(y)$, а именно решения вида

$$\Psi(x, y) = \sum_i u_i(x)\, v_i(y). \tag{14}$$

Чему соответствуют такие решения и когда они встречаются? Это очень важный класс решений: с ними приходится иметь дело почти всякий раз, когда происходит измерение.

Предположим, что мы заставляем две первоначально не связанные системы I и II взаимодействовать в течение короткого времени (например, когда частицы сталкиваются) и затем вновь их развязываем. Краткость взаимодействия может проистекать, например, из того, что оно имеет место лишь на малых расстояниях. Таким образом, мы имеем первоначально функцию (12), затем на короткое время уравнение для Ψ, содержащее член

$V_{12}(x, y)$, который описывает взаимодействие

$$\frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2} - \frac{2m}{\hbar^2}[V_1(x) + V_2(y) + V_{12}(x, y)]\Psi = -\frac{2mi}{\hbar}\frac{\partial\Psi}{\partial t}. \quad (15)$$

Для него начальное выражение Ψ задается по (12), и из него мы получаем выражение для $\Psi(x, y)$, которое, в свою очередь, является начальным для третьей стадии, когда системы вновь не взаимодействуют друг с другом. Так вот, в этой третьей стадии мы и имеем не произведение, а сумму произведений (14).

Естественно возникает вопрос: если известна функция $\Psi(x, y)$ (я ее могу рассчитать), то какая же функция $\psi(x)$ характеризует систему *I*, когда она уже сильно отошла от системы *II*? Какая функция характеризует поведение электрона, а какая — поведение α-частицы? Нейман первый обратил внимание на то, *что такой функции нет.* Но ведь системы уже разошлись, ведь есть же какое-то поведение у системы *I*?

Ответ на этот недоуменный вопрос заключается в следующем. Можно ответить на все вопросы, касающиеся системы *I*, т. е. (при любом состоянии *II*) найти вероятности значений различных величин для этой системы (вероятности различных значений импульса и т. д.), но функции $\psi(x)$, которая характеризовала бы состояние системы *I* вне зависимости от состояния системы *II*, все же нет.

Прежде всего удостоверимся, что здесь нет ничего особенно поразительного. Действительно, если существует функция $\psi(x)$, то мы знаем, как следует связывать между собой статистики различных величин: надо разлагать эту $\psi(x)$ по собственным функциям операторов рассматриваемых величин. Тем самым статистики этих величин определенным образом связаны между собой через $\psi(x)$. Но откуда следует, что *для системы после взаимодействия* статистики различных физических величин будут связаны между собой так же, как и при наличии $\psi(x)$? Это ниоткуда не следует, и, как правило, этого как раз нет.

Нетрудно убедиться на простом примере, что случаи, когда $\psi(x)$ не существует, действительно могут иметь место. Пусть функция для совокупности систем *I* и *II* есть

$$\Psi(x, y) = a_1\psi_1(x)\varphi_1(y) + a_2\psi_2(x)\varphi_2(y) + a_3\psi_3(x)\varphi_3(y). \quad (16)$$

Здесь ψ_i — собственные функции какого-то оператора R, φ_i — оператора S. Что можно сказать о системе *I* на основании формулы (16)? Очевидно, для нее отличны от нуля вероятности состояний 1, 2 и 3, но равны нулю вероятности состояний 4, 5, ... Следовательно, если существует $\psi(x)$ для системы *I*, то она должна

иметь вид

$$\psi(x) = c_1\psi_1(x) + c_2\psi_2(x) + c_3\psi_3(x). \tag{17}$$

Теперь нужно потребовать, чтобы статистика любой величины, относящейся к системе I, получалась одной и той же из (16) и из (17). Но, например, вероятность координаты x (независимо от того, каково y) из (16) получается равной

$$\int \Psi\overline{\Psi} dy = a_1\bar{a}_1\psi_1\overline{\psi}_1 + a_2\bar{a}_2\psi_2\overline{\psi}_2 + a_3\bar{a}_3\psi_3\overline{\psi}_3,$$

а из (17)

$$\psi\overline{\psi} = c_1\bar{c}_1\psi_1\overline{\psi}_1 + c_2\bar{c}_2\psi_2\overline{\psi}_2 + c_3\bar{c}_3\psi_3\overline{\psi}_3 + c_1\bar{c}_2\psi_1\overline{\psi}_2 + \ldots$$

Не представляет труда выбрать такой случай, когда $\psi_1\overline{\psi}_1, \psi_2\overline{\psi}_2, \psi_3\overline{\psi}_3, \psi_1\overline{\psi}_2, \ldots$ линейно независимы. Значит, для совпадения результатов необходимо, чтобы было $c_1\bar{c}_1 = a_1\bar{a}_1$, $c_2\bar{c}_2 = a_2\bar{a}_2$, $c_3\bar{c}_3 = a_3\bar{a}_3$ и вместе с тем $c_1\bar{c}_2 = \bar{c}_1c_2 = c_2\bar{c}_3 = \bar{c}_2c_3 = c_3\bar{c}_1 = \bar{c}_3c_1 = 0$. Но это невозможно. Следовательно, нельзя подобрать такие c_1, c_2 и c_3, при которых (17) давало бы такую же статистику, как и (16), т. е. функции $\psi(x)$ для системы I не существует [1]. Физически в этом нет ничего невероятного: статистики для различных величин у системы I есть, но они связаны между собой не так, как при наличии $\psi(x)$.

Итак, встречаются случаи, не описываемые функцией $\psi(x)$ и наступающие обычно после взаимодействия. Они называются *смесями* в отличие от *чистых случаев* (чистых состояний), описываемых функцией $\psi(x)$. Для объединенной системы существует $\Psi(x, y)$, но не существует в отдельности взятых функций $\psi(x)$ и $\varphi(y)$.

Спрашивается, нельзя ли выбрать такую подсовокупность, для которой функция $\psi(x)$ существует? Очевидно, это можно сделать, если для системы I мы будем принимать во внимание только такие состояния, когда состояние системы II выделено (определенным значением либо координаты, либо какой-нибудь другой величины). Например, если мы в статистике, описывающей электрон, будем принимать во внимание только те случаи, когда α-частица имеет определенную координату, то легко ви-

---

[1] [Если ψ_i вещественны, как, например, для осциллятора, то

$$\psi\overline{\psi} = c_1\bar{c}_1\psi_1^2 + c_2\bar{c}_2\psi_2^2 + c_3\bar{c}_3\psi_3^2 + (c_1\bar{c}_2 + c_2\bar{c}_1)\psi_1\psi_2 + \ldots$$

и должны одновременно соблюдаться условия

$$\frac{\bar{c}_1}{c_1} = -\frac{\bar{c}_2}{c_2} = \frac{\bar{c}_3}{c_3} = -\frac{\bar{c}_1}{c_1},$$

что невозможно если все c_i не равны нулю.]

деть, что для такой подсовокупности поведение электрона описывается ψ-функцией.

После этих общих замечаний вернемся теперь к вопросу о косвенных измерениях.

Пусть мы хотим измерить в системе I (координата x) некоторую величину λ (оператор R с собственными значениями λ_1, λ_2, ... и собственными функциями $\psi_1(x)$, $\psi_2(x)$, ...). Мы приводим систему I во взаимодействие с системой II (координата y), в которой *после взаимодействия* мы можем осуществить *прямое измерение* некоторой величины μ (оператор S с собственными значениями μ_1, μ_2, ... и собственными функциями $\varphi_1(y)$, $\varphi_2(y)$, ...). Если я утверждаю, что такое устройство может служить для определения λ_i по результатам прямого измерения μ_i, то это предполагает, что между μ_i и λ_i существует однозначное (или достаточно полное) соответствие.

Самый известный пример такого типа измерения, описываемый во всех учебниках, это гейзенберговский микроскоп: если мы хотим определить положение какой-либо частицы, то мы можем «осветить» ее потоком других частиц, например электронов, и по положению пятен, создаваемых этими частицами на фотографической пластинке, судить о координате интересующей нас частицы. Разумеется, если система II — фотоны, то она будет описываться не шрёдингеровским уравнением. Но это не принципиальный момент, микроскоп может быть и электронным, ибо нас интересуют вопросы, не зависящие от частных особенностей фотона. При этом мы будем считать, что мы находимся в нерелятивистской области.

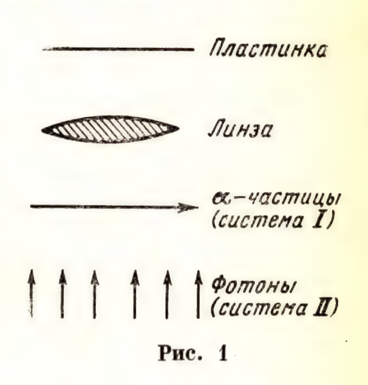


Рис. 1

До соударения мы имеем $\psi_0(x)$ для системы I и $\varphi_0(y)$ для системы II, т. е.

$$\Psi_0(x, y) = \psi_0(x)\,\varphi_0(y). \qquad (18)$$

После соударения, как правило, мы имеем не произведение, а смесь [1], но мы всегда можем разложить $\Psi(x, y)$ по собственным функциям оператора S, относящегося ко второй системе

$$\Psi(x, y) = \sum_i u_i(x)\,\varphi_i(y). \qquad (19)$$

Входящие сюда коэффициентами функции $u_i(x)$ не обязательно ортогональны. Предположим теперь, что функция $\psi_0(x)$ была

[1] Для всей системы у нас чистый случай, но для каждой подсистемы — смесь.

собственной функцией нашего оператора R, скажем, что $\psi_0 = \psi_i$, т. е. до взаимодействия было $\lambda = \lambda_i$. Вообще говоря, $\Psi (x, y)$ будет смесью, но может случиться, что для рассматриваемого процесса взаимодействия, при $\psi_0 (x) = \psi_i (x)$, после взаимодействия получается простое произведение двух функций, из которых функция, описывающая вторую систему $\varphi_i (y)$, принадлежит собственному значению $\mu = \mu_i$ оператора S, одно-однозначным образом связанному с собственным значением λ_i оператора R, которым система I обладала до взаимодействия

$$\Psi = u_i (x) \varphi_i (y). \tag{20}$$

Тогда, зная, что в системе II после взаимодействия $\mu = \mu_i$, мы тем самым знаем, что в системе I до взаимодействия было $\lambda = \lambda_i$ [1]. Если такое устройство, в котором для всех i выполняется (20), найдено, то мы скажем, что мы нашли аппарат для измерения значения величины R, имевшего место до измерения [2].

Волновая механика постулирует, что такое устройство можно найти для всякой физической величины R (*хотя бы в каком-то n-м звене*). Без такого постулата значения λ теряют физический смысл. Обладает ли некоторое данное устройство таким свойством, это можно просто вычислить. Разумеется, устройство, являющееся измерителем для некоторой величины, может не быть измерителем для какой-либо другой величины, т. е. для нее уже не будет справедливо (20).

Вот, в сущности, вся основа косвенных измерений в волновой механике. Эти положения имеют далеко идущие следствия, к которым мы теперь и перейдем.

---

[1] Совершенно очевидно, что можно требовать гораздо меньше: если λ_i дает, скажем, три значения μ, но каждое λ_i — свои три значения μ, то этого достаточно для определения λ по μ. Однако для простоты мы остановимся на одно-однозначном соответствии.

[2] [Когда две системы I и II приводятся во взаимодействие в измерительном устройстве, то, вообще говоря, всегда существует конечная вероятность того, что состояния обеих систем после этого останутся неизменными. Например, когда освещают некоторую частицу I для того, чтобы определить ее положение, то фотон (система II) может пройти мимо частицы, не испытав рассеяния или поглощения. Таким образом, конечное состояние системы $I + II$ описывается наложением начальной функции системы $\Psi_0 (x, y)$ («проходящая волна») и видоизмененной взаимодействием функции $\Psi' (x, y)$ («рассеянная волна»).

В дальнейшем предполагается, что измерительное устройство регистрирует только рассеянную волну, что всегда осуществимо с желаемой степенью точности. Если же состояние системы после «прохождения» через измерительное устройство не изменилось («проходящая волна»), то считается, что как взаимодействие систем, так и самое измерение не осуществилось. Соответственно этому утверждение, что если начальное состояние системы I описывается функцией $\psi_i (x)$, то после взаимодействия систем I и II в измерительном устройстве получается состояние, описываемое функцией $u_i (x) \varphi_i (y)$, относится к «рассеянной волне».]

Если величину λ в системе I можно измерить прямо и, кроме того, косвенно, то, поскольку во втором случае привходит теория — уравнение Шрёдингера, постольку мы имеем здесь *проверку теории.* Конечно, если прямого измерения для λ нет, то косвенное измерение ничего не проверяет, а является, по сути дела, *определением* λ. Но может случиться, что прямого измерения нет, зато есть несколько косвенных. Тогда мы опять имеем проверку теории. Эти последние соображения применимы, конечно, и в классике. Например, измерение положения «материальной точки» может быть сделано и прямым измерением (совпадение с делением масштаба) и косвенным (например, при помощи микроскопа). Совпадение обоих значений есть проверка оптических законов. Если бы прямое измерение было невозможно, то измерение координаты микроскопом должно было бы служить ее определением. Но тогда, конечно, о какой-нибудь проверке оптики нет речи. Но если я возьму два или несколько различных оптических измерений и результаты совпадут, то, хотя прямое измерение невозможно, совпадение отдельных результатов служит проверкой оптики.

Мы рассмотрели случай, когда начальная $\psi(x)$ была собственной функцией: $\psi_0(x) = \psi_i(x)$. Если этого нет, то я всегда могу разложить $\psi_0(x)$ по собственным функциям

$$\psi_0(x) = \sum_i c_i \psi_i(x) \tag{21}$$

и, зная, что́ дает каждая $\psi_i(x)$ после взаимодействия, записать результирующую $\Psi(x, y)$. Очевидно, в силу линейности уравнения Шрёдингера

$$\Psi(x, y) = \sum_i c_i u_i(x)\, \varphi_i(y). \tag{22}$$

При этом, вообще говоря, $u_i(x)$ не равно $\psi_i(x)$.

Величины μ_i, соответствующие функциям $\varphi_i(y)$, мы считаем измеряемыми прямо. Эти прямые измерения μ после взаимодействия дадут значение μ_i с вероятностью $c_i\bar{c}_i$ (при любых конечных состояниях системы I). Поэтому имеет непосредственное содержание высказывание, что $c_i\bar{c}_i$ есть вероятность получения значения μ_i, а следовательно, и λ_i ввиду однозначного соответствия между ними. Если бы в системе I существовал прямой метод измерения λ_i, то, согласно (21), вероятность значения λ_i была бы $c_i\bar{c}_i$. Отсюда видно, что статистика (выражение для вероятности), получаемая косвенным измерением, та же, что и получаемая прямым [из $\psi_0(x)$], если прямое измерение возможно. Если прямое измере-

ние невозможно, то статистика для μ_i служит определением статистики для λ_i.

В первом случае я имею существенный результат: статистика инвариантна по отношению к тому этапу, где мы переходим на классику (именно, она одинакова при прямом измерении, где классика уже в первом этапе, и в косвенном, где классика во втором этапе). Как видите, это очень существенное обстоятельство, показывающее, что во всей постановке вопроса нет внутреннего противоречия.

Далее, $\Psi_0(x, y)$ зависит не только от того, каково $\psi_0(x)$, но и от $\varphi_0(y)$. При одних $\varphi_0(y)$ данное устройство может являться измерителем λ, при других не может. Здесь особенно ясно видно, что в волновой механике само определение величины зависит от прибора, которым вы работаете. Если это постоянно иметь в виду, то можно избегнуть многих «парадоксов». Вот, например, один из них, на который я сейчас только укажу: собственные функции осциллятора $\psi_n(x)$ стремятся к нулю при $x \to \infty$, но для всякого сколь угодно большого *конечного* x: $|\psi(x)|^2 \neq 0$. Таким образом, существует отличная от нуля вероятность того, что потенциальная энергия осциллятора $V = \frac{\alpha x^2}{2}$ *превышает* его полную энергию $\varepsilon_n = \left(n + \frac{1}{2}\right)h\nu$. Это, говорят, очень неприятно, это парадокс.

Прежде всего, я не вижу, почему это шокирует. В классике мы знаем, что $E = V + T$, причем $T = \frac{mv^2}{2} > 0$. Здесь действительно неприятно, если $V > E$. Но в волновой механике мы ведь вообще не можем измерять одновременно x и v, а стало быть, и V и T.

Однако если здесь и нет парадокса, то может быть существенное отличие от классики. Рассмотрим вопрос об измерении таких больших x, что $V > E$. Можно ли констатировать (измерить), что осциллирующая частица претерпевает столь большие отклонения от положения равновесия? Оказывается, как мы увидим далее, что не всякий измеритель x пригоден для измерения таких отклонений, при которых $V > E$. Именно поэтому, пока мы рассуждаем оторванно от измерителя, и получается парадокс. Необходимо рассматривать измеритель и измеряемую величину в совокупности: измерительное устройство, т. е. физические условия, в которых находится частица, существенно определяет возможные значения измеряемой величины. В данном случае обычные измерительные устройства, годные при малых x, когда $V < E$, вообще говоря, не будут годны для измерения положения частицы вдали от положения равновесия, когда $V > E$, если

только эти устройства не могут при измерении передавать измеряемой частице достаточно большой энергии [1].

Я уже указывал, что нужно каждый раз изыскивать определенное измерительное устройство. Но встречалась и даже превалировала другая точка зрения. Так, в первом издании книги Дирака (во втором издании этого вообще нет) утверждалось, что два измерения, очень быстро следующих одно за другим, дают одно и то же («повторяемость» значений). Так ли это?

Пусть перед первым измерением для системы I мы имели $\psi_1(x)$. Тогда после взаимодействия получим $\Psi(x, y) = u_1(x)\varphi_1(y)$ [2]. Таким образом, для второго измерения мы имеем начальную функцию $u_1(x)$, вообще говоря, не равную $\psi_1(x)$. Значит, и $\Psi(x, y)$ после второго измерения будет другим. Иными словами, нигде не сказано, что измерительное устройство обязательно позволяет повторять измерение с теми же результатами. Это может иметь место, но, вообще говоря, лишь в исключительных случаях. Например, измерения координат повторимы [3], а импульса — нет.

Остановимся в заключение еще на одном интересном вопросе, на так называемой редукции волнового пакета. Обычное изложение этого вопроса мне представляется какой-то мистикой. Между тем дело, по-моему, совсем просто.

Как ставился вопрос?

Имеется пакет, изображающий некоторую частицу и падающий на потенциальный барьер. Пусть барьер таков, что пакет наполовину отражается и наполовину проходит. Мы получаем,

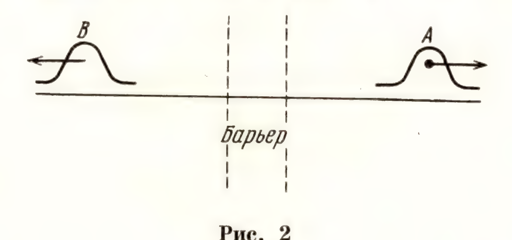


Рис. 2

следовательно, два пакета, разбегающихся в противоположные стороны (рис. 2). Мы подождем, пока A и B не удалятся друг от друга чрезвычайно далеко, и затем произведем измерение, скажем, с помощью микроскопа.

Пусть в результате этого измерения мы нашли, что частица находится в A. Тогда после измерения, если есть волновая функ-

[1] [Этот вопрос детально разбирается ниже, в 4-й лекции.]
[2] [См. примечание 2 на стр. 344.]
[3] [См. стр. 360.]

ция для изучаемой частицы, то она отлична от нуля только в A, а в B ничего нет. Говорят, что акт измерения положения частицы редуцирует волновой пакет (собирает его в A). По этому поводу Эйнштейн сделал на Сольвеевском конгрессе замечание, смысла которого я, сколько ни думал, не понимаю. Действие в A, сказал Эйнштейн, изменило вероятность в B от $1/2$ до 0. Так как расстояние между A и B очень велико, а действие в A могло быть очень кратким, то, следовательно, действие в A передалось в B со скоростью, которая может в любое число раз превышать скорость света. Эйнштейну возразили, что это действие не является сигналом, и на этом успокоились.

По-моему, дело вообще совсем в другом и совершенно не относится к квантовой механике. Здесь попросту нет никакого *действия*. Разве вероятность есть что-то такое, на что физически действуют? До измерения (с помощью системы II — координата y) мы имели $\Psi_0(x, y) = \psi_0(x)\varphi_0(y)$; $\psi_0(x)$ — это и был (уже раздвоившийся) пакет. После измерения мы получаем $\Psi(x, y) = \sum_i c_i u_i(x)\varphi_i(y)$, т. е. теперь вообще нет волновой функции $\psi(x)$ и нет никакого пакета. Можно, однако, выделить подсовокупность, т. е. выбрать только те случаи, когда прямое измерение над системой II дает, скажем, $\mu = \mu_1$. В пределах этой подсовокупности для системы I существует волновая функция, а именно, $u_1(x)$. Это выделение волновой функции и чистого состояния происходит за счет добавочного классического измерения, относящегося к y. Это и есть «редукция волнового пакета», образование нового пакета. Ничего удивительного в том, что этот пакет отличен от прежнего, разумеется, нет.

Предположите, что некое лицо N ездит из Ленинграда в Москву и обратно, проводя в среднем одинаковое время в каждом из этих городов. В Москве и Ленинграде сидят наблюдатели и производят, скажем, ежедневно, «измерения», а именно, констатируют наличие или отсутствие N. Длительная статистика покажет им, что вероятность нахождения в каждом из городов равна $1/2$. Но если мы примем в расчет только те наблюдения ленинградского наблюдателя, при которых произведенные в тот же день наблюдения москвича дали 1, то мы получим, что вероятность пребывания в Ленинграде равна нулю. Вот и вся «редукция волнового пакета». Очевидно, здесь нет никакого противоречия с теорией относительности. Если москвич поставит себе задачей предупредить ленинградца о своем результате, то он будет телеграфировать. Телеграмма — это, конечно, «действие», это сигнал, но она и идет не быстрее света в вакууме.

Сказанное интересно связать с самым первым нашим положением о том, что совокупность задается макроскопически. До

взаимодействия мы задаем некоторое количество макроскопических параметров, достаточное для того, чтобы определить волновые функции $\psi(x)$ и $\varphi(y)$ систем *I* и *II*. После взаимодействия этих параметров обычно уже недостаточно, и для определения парциальных волновых функций надо задать некоторые макропараметры дополнительно. Этим и выделяется подсовокупность, в пределах которой снова существует $\psi(x)$.

ТРЕТЬЯ ЛЕКЦИЯ

(20.IV 1939 г.)

Резюмируем кратко сформулированные нами общие положения квантовомеханической теории измерений.

Чтобы косвенно измерить относящуюся к системе *I* величину R с собственными значениями λ, мы приводим систему *I* во взаимодействие с другой, тоже микроскопической системой *II*, в которой мы умеем прямо, классическими опытами, измерять величину S, обладающую собственными значениями μ_k. До взаимодействия имеем

$$\Psi_0(x, y) = \psi_0(x)\,\varphi_0(y). \tag{23}$$

После взаимодействия [1] всегда можно представить Ψ в виде разложения по собственным функциям оператора S

$$\Psi(x, y) = \sum_i c_i u_i(x)\,\varphi_i(y). \tag{24}$$

Коэффициенты c_i мы теперь ввели для того, чтобы $\varphi_i(y)$ можно было считать нормированными.

Если условия таковы, что при $\psi_0(x) = \psi_i(x)$ после взаимодействия будет

$$\Psi(x, y) = u_i(x)\,\varphi_i(y), \tag{25}$$

то, измеряя состояние системы *II* после взаимодействия, мы можем заключить о состоянии системы *I* до взаимодействия. Требование, чтобы такое соотношение выполнялось, есть требование того, чтобы данная аппаратура являлась измерителем величины λ.

---

[1] [См. примечание 2 на стр. 344.]

После измерения мы, вообще говоря, имеем смесь. Но мы можем выделить подсовокупность (скажем, тем, что выберем определенное $\mu = \mu_k$), в которой для системы I существует волновая функция, а именно, $u_k(x)$. В общем случае $u_k(x) \neq \psi_0(x)$. Возможно, что система II имеет не одну, а две степени свободы — практика обычно гораздо сложнее, чем удобные для дискуссии примеры. Тогда

$$\Psi(x, y, z) = \sum_{i,j} c_{ij} u_{ij}(x) \varphi_i(y) \chi_j(z). \tag{26}$$

Одновременно, измеряя собственные значения μ_i и ν_j для систем II и III, мы выделяем подсовокупность, в которой система I описывается волновой функцией $u_{ij}(x)$.

Для непрерывных спектров суммы заменятся интегралами.

Все эти общие положения несколько абстрактны. Теперь мы проиллюстрируем их на примерах измерения двух основных величин — импульса и координаты. Дискуссия этих примеров представляет интерес, с одной стороны, потому, что в них мы уже будем иметь дело с определенными физическими измерительными устройствами (конечно, упрощенными), с другой стороны, мы рассмотрим действие этих устройств с точки зрения нашей схемы.

Прежде всего рассмотрим *измерение импульса свободной частицы*.

Если на частицу не действуют силы и ее энергия есть E, то решением уравнения Шрёдингера является функция

$$\psi'(x, t) = e^{\frac{i}{\hbar}\sqrt{2mE}\, x} \cdot e^{-\frac{i}{\hbar}Et} \tag{27}$$

(штрих отмечает, что это решение временно́го уравнения). Эта функция является собственной функцией оператора импульса. Следовательно, отбрасывая временно́й множитель, ее можно записать так:

$$\psi(x) = e^{\frac{i}{\hbar}p'x}.$$

Отсюда получается связь между энергией и импульсом

$$p' = \sqrt{2mE}. \tag{28}$$

Я намеренно ничего не говорю пока о *скорости*. Мы пока ввели определенным образом только координату и импульс (или, лучше сказать, квазикоординату и квазиимпульс).

Если частица заряженная (протон или α-частица), то ее импульс можно измерить чисто классически — прямым опытом, например по ее пробегу в камере Вильсона. Но для нейтрона этот путь закрыт, и нам нужно найти устройство, позволяющее изме-

рять импульс нейтрона косвенно. Его можно измерять, например, по пробегу протона, с которым мы заставляем этот нейтрон соударяться. Для простоты допустим, что все происходит в одном измерении. (Фактически осуществить это нельзя, но все важные нам принципиальные моменты при таком рассмотрении сохраняются.)

Пусть нейтрон падает на пластинку водородсодержащего вещества (например, парафина) и выбивает из него протон, который влетает в стоящую за пластинкой камеру Вильсона (рис. 3). По пробегу этого протона в камере Вильсона, поддающемуся прямому измерению, мы и хотим узнать импульс нейтрона.

Для падающего на пластинку нейтрона (системы I) мы имеем [1]

$$\psi_0(x) = e^{\frac{i}{\hbar} px}. \tag{29}$$

Протон заключен в толще пластинки, и если она велика по сравнению с длиной волны, то можно считать, что импульс p_1 у протона какой-то определенный. Именно, можно положить

$$\varphi_0(y) = f(y)\, e^{\frac{i}{\hbar} p_1 y}, \tag{30}$$

где $f(y)$ — медленно меняющаяся функция, не равная нулю

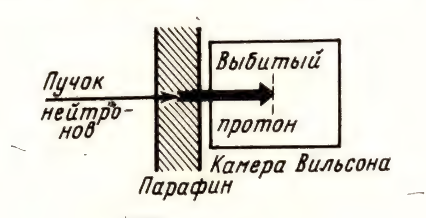


Рис. 3

внутри пластинки и исчезающая вне ее. Таким образом, она мало меняется на протяжении одной длины волны. Протон прак-

[1] Я хочу сделать одно общее замечание, не специфичное для волновой механики. Его, однако, следует иметь в виду при переходе к обсуждению опытов. Мы говорим об определенном импульсе, когда $\psi = \exp\left(\frac{i}{\hbar} px\right)$, причем эта формула должна быть верна во всем пространстве ($-\infty < x < +\infty$). Фактически такую волну получить нет возможности. Но это то же самое, что в оптике с монохроматическим светом. И там замена достаточно монохроматического света (достаточно длинных цугов) идеальным монохроматическим для многих вопросов (конечно, не при рассмотрении групповой скорости) представляет собой вполне законную идеализацию.

тически не движется, так что можно считать $p_1 = 0$. Следовательно,

$$\varphi_0(y) = f(y) \tag{31}$$

— медленно меняющаяся функция. Итак, до соударения

$$\Psi_0(x, y) = f(y)\, e^{\frac{i}{\hbar} px}. \tag{32}$$

Условие, чтобы функция $\psi_0(x)$ являлась собственной функцией оператора измеряемой величины (в нашем случае — импульса), здесь выполнено. Будет ли наше устройство измерителем импульса? Для этого нужно посмотреть, каково $\Psi(x, y)$ после взаимодействия. Если наше устройство может служить измерителем импульса, то после взаимодействия должно быть

$$\Psi(x, y) = u(x)\varphi(y).$$

Фактически мы не включаем и не выключаем взаимодействия, но последовательность во времени обусловливается перемещением частиц. При этом взаимодействие весьма кратковременно, так как оно имеет место лишь на малых расстояниях. Мы должны написать теперь для Ψ уравнение Шрёдингера и учесть в нем это взаимодействие добавочным членом.

Мы упростим задачу, положив, что взаимодействие является отталкиванием, что оно бесконечно велико при совпадении частиц ($x = y$) и равно нулю во всех остальных точках.

Что означает бесконечно большое отталкивание при $x = y$? Это значит, что вероятность найти там что-нибудь равна нулю [1]:

$$\Psi(x, y) = 0 \text{ при } x = y. \tag{33}$$

С этим условием мы должны обратиться теперь к уравнению для Ψ, но в самом уравнении никаких членов, описывающих взаимодействие, учитывать уже не надо. Так как массы у нейтрона и протона, к счастью, одинаковы, то уравнение будет

$$\frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2} + \frac{p^2}{2m\hbar^2}\Psi = 0, \tag{34}$$

где $p^2/2m$ — общая энергия системы. Разумеется, функция (32) удовлетворяет этому уравнению, если только помнить, что $f(y)$ — почти постоянная величина в широком интервале.

---

[1] [Условие (33) можно вывести, рассматривая модель непроницаемых частиц исчезающе малого размера в одном измерении. В трех измерениях условие для Ψ формулируется на поверхности сферы конечного радиуса; см. уравнение (66).]

Вместо того чтобы решать формулированную выше задачу, я сошлюсь теперь на точно такую же задачу из оптики — способ, который во многих случаях оказывается полезным. Именно, возьмем действительно двухмерную задачу, когда волновое уравнение

$$\frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + k^2\Phi = 0$$

такое же, как (34), возьмем *широкий пучок света*, идущий в направлении x

$$\Phi_0 = f(y)\, e^{ikx},$$

т. е. зададим его той же формулой, что и (32), и, наконец, поставим граничное условие (33)

$$\Phi\,|_{x=y} = 0.$$

Очевидно, при такой постановке вопроса мы имеем задачу об отражении от идеального зеркала, поставленного под углом 45°

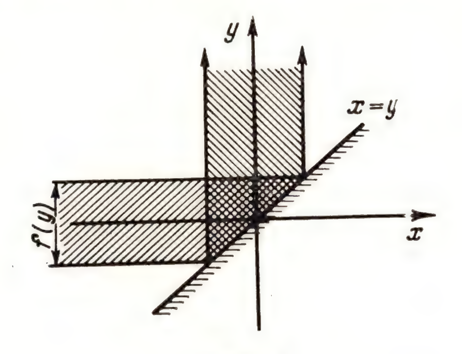


Рис. 4

к падающему пучку (рис. 4). Результат известен: получится такой же пучок, идущий в направлении оси y, т. е. $\Psi(x, y)$ после взаимодействия есть

$$\Psi_1(x, y) = -f(x)\, e^{\frac{i}{\hbar}py}. \qquad (35)$$

Это значит, что теперь нейтрон покоится [нейтроны застряли в парафине, и собственная функция системы I изменилась $e^{\frac{i}{\hbar}px} \to f(x)$], а протон имеет импульс p, измерять который мы умеем прямо.

Функция (35) имеет как раз такой вид, который, согласно (25), необходим, чтобы по состоянию системы II после взаимодействия можно было заключить о состоянии системы I до взаимодействия, т. е. наше устройство действительно является измерителем импульса. Если $\mu_i = p$, то мы заключаем отсюда, что до взаимодействия нейтрон обладал импульсом $\lambda_i = p$.

Как видно из этого примера, введенные нами величины p и E обладают тем свойством, что их суммарное значение сохраняется в процессе взаимодействия. Это чрезвычайно важно, потому что как раз таким свойством обладают классические импульс и энергия. Поэтому при нашем определении импульса и энергии в квантовой механике получаются те же законы сохранения, что и в классике. Между тем законы сохранения вообще и есть то, что придает в механике смысл введению понятий энергии и импульса.

С другой стороны, если для нейтрона нет других способов измерения p, то наш опыт как раз и представляет собой определение понятия «импульс нейтрона». Тем не менее наш опыт имеет большее содержание, чем простое определение. Дело в том, что можно проделать такой же опыт, воспользовавшись частицей с иной массой, чем у протона. Тогда, сравнивая результаты разных опытов, мы уже будем иметь проверку теории. Такой опыт с частицей другой массы сведется (в оптической аналогии) к тому, что зеркало будет стоять под углом, отличным от 45°. В самом деле, если масса второй частицы не m, а αm, то в (34) перед членом $\frac{\partial^2\Psi}{\partial y^2}$ появится коэффициент $1/\alpha$. Поэтому, рассматривая все в системе координат не (x, y), а (x, y'), где $y' = \sqrt{\alpha}y$, мы из условия $\Psi = 0$ при $x = y$ получим условие $\Psi = 0$ при $x = \frac{1}{\sqrt{\alpha}}y'$. Если масса рассеивающей частицы очень велика ($\alpha \gg 1$), то $\Psi = 0$ на всей плоскости $x = 0$, зеркало станет поперек пучка и вернет нейтрон обратно (нейтрон отразится). Законы сохранения отсюда легко вывести, и мы получаем возможность проверять теорию [1].

[1] Необходимо сделать одно замечание: как до удара, так и после удара мы имеем произведение функций от x и y. Это как раз требуется для возможности производить измерения. Но в дважды заштрихованном на рис. 4 пространстве у нас

$$\Psi(x, y) = f(y)\, e^{\frac{i}{\hbar}px} - f(x)\, e^{\frac{i}{\hbar}py},$$

т. е. здесь нет простого произведения. Здесь, таким образом, нельзя «измерять». «Время пробега» частицы в этой области $\Delta t \sim l/v$, где l — ее ширина. Однако, применяя оптическую аналогию, мы пользовались геометрической оптикой, т. е. пренебрегли дифракцией, считая $f(y)$ функцией, достаточно медленно меняющейся по длине волны, т. е. считая, что $l \gg \hbar/p$. Это означает, что производить измерение можно, только подождав время $\Delta t \gg \hbar/mv^2$ после удара.

Перейдем теперь ко второму примеру — *к измерению координаты*. Мы воспользуемся для этого электронным микроскопом, с помощью которого всегда излагают эти вопросы, причем почти всюду, если не всюду, излагают неправильно.

Следует отметить, что и здесь мы будем касаться лишь вопроса о *точном* измерении. В предыдущем примере мы показали, как можно *точно* измерить импульс. Немонохроматичность у нас была, но не принципиальная: она возникла из-за конечной толщины пластинки парафина. В принципе же мы имели $\Delta q = \infty$, и поэтому, не нарушая соотношения $\Delta p \cdot \Delta q \geqslant \hbar$, у нас могло быть $\Delta p = 0$. О неточности лучше говорить не сейчас, а далее, при дискуссии приближенных измерений. Теперь же, как и в примере с импульсом, мы рассмотрим задачу о точном (когда это возможно) косвенном измерении координаты.

Итак, мы имеем следующее устройство: частица M, координату которой мы хотим измерить, может двигаться только вдоль оси z (рис. 5); по направлению оси x падает пучок электронов, имеющих массу m, причем координаты электрона в плоскости чертежа обозначим x и y (y — координата электрона по оси z). Рассеянные электроны собираются линзой на пластинку, где и дают изображение. Может ли такой микроскоп измерять значения координаты частицы?

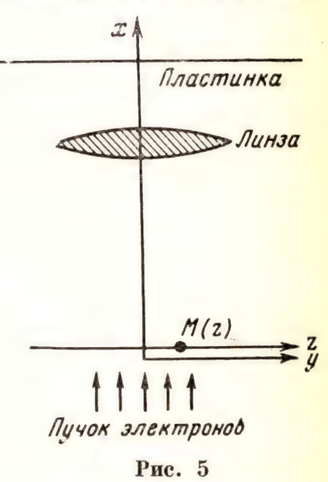


Рис. 5

Для упрощения мы заменим частицу дырочкой (дополнительным экраном, по Бабине), что позволит нам рассматривать только поле, рассеянное «частицей». Без такой замены пришлось бы учитывать и первичное поле. Замена частицы движущейся диафрагмой (с той же массой) может показаться странной, но очевидно, что она возможна и законна. Мы имеем уравнение

$$\frac{1}{2m}\left(\frac{\partial^2\Psi}{\partial x^2}+\frac{\partial^2\Psi}{\partial y^2}\right)+\frac{1}{2M}\frac{\partial^2\Psi}{\partial z^2}+\frac{E}{\hbar^2}\Psi=0, \tag{36}$$

причем до соударения, когда частицы не взаимодействуют,

$$\Psi_0(x, y, z)=e^{\frac{i}{\hbar}qx}e^{\frac{i}{\hbar}pz}, \qquad \left(E=\frac{q^2}{2m}+\frac{p^2}{2M}\right), \tag{37}$$

где q — импульс электрона по оси x, p — импульс частицы. Как

учесть, что электроны могут проходить только через бесконечно малую дырочку? Мы должны, очевидно, выразить то обстоятельство, что прохождение имеет место лишь при совпадении координаты электрона y с координатой «частицы» (на плоскости $x = 0$). Очевидно, мы достигнем цели, приняв, что при $x = 0$

$$\Psi = e^{\frac{i}{\hbar} pz} \delta(z - y). \tag{38}$$

Легко построить оптическую аналогию нашей задачи, введя y как настоящую третью координату, ортогональную к x и z. Это будет задача о дифракции на щели.

Если воспользоваться равенством

$$\int_{-\infty}^{+\infty} e^{-\frac{i}{\hbar} \alpha r} dr = 2\pi\hbar\delta(\alpha), \tag{39}$$

то условие (38) можно переписать в виде

$$\Psi(0, y, z) = e^{\frac{i}{\hbar} pz} \delta(z - y) = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} e^{\frac{i}{\hbar} [ry + (p-r) z]} dr. \tag{40}$$

Итак, нам нужно написать решение уравнения (36), соответствующее заданной энергии E и удовлетворяющее условию (40), которое выражает мгновенный акт взаимодействия электрона с частицей-дырочкой. Нетрудно убедиться простой подстановкой, что искомое решение есть

$$\Psi(x, y, z) = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} e^{\frac{i}{\hbar} [ry + (p-r) z + \sqrt{k^2 - \varepsilon^2 (p-r)^2 - r^2}\, x]} dr, \tag{41}$$

где

$$\varepsilon^2 = \frac{m}{M}, \qquad k^2 = 2mE. \tag{42}$$

Что это решение означает? Оно представляет собой интеграл (сумму) произведений трех экспонентов, т. е. трех собственных функций операторов проекций импульса. Коэффициент при каждом слагаемом по общему правилу дает вероятность того, что компоненты импульса по осям y, z и x будут соответственно равны

$$r,\ p - r \ \text{ и } \sqrt{k^2 - \varepsilon^2 (p - r)^2 - r^2}\,.$$

Таким образом, в смесь входят лишь такие члены, для которых выполняются вполне определенные соотношения между импульсами

$$p_y + p_z = r + (p - r) = p = \text{const}, \tag{43}$$

$$\frac{p_x^2 + p_y^2}{2m} + \frac{p_z^2}{2M} = E. \tag{44}$$

Импульс по направлению совпадающих друг с другом осей y и z сохраняется, причем очевидно, что электрон приобретает импульс r по оси z за счет импульса частицы ($p \to p - r$), т. е. здесь сказывается обратное действие измерителя на измеряемое. Это, по существу, и соответствует комптон-эффекту. По оси x сохранения импульса нет: до взаимодействия импульс был q, а теперь он не равен q. Конечно, смысл этого не тот, что нарушен закон сохранения, а тот, что мы наложили связь, запретив диафрагме (частице) двигаться по оси x. Другими словами, импульс по оси x у частицы есть, но мы его не учитываем. Отсюда и проистекает несохранение импульса p_x. Наконец, из (37) и (44) мы видим, что наше решение удовлетворяет сохранению энергии.

Нам нужно теперь, для того чтобы микроскоп был измерителем координаты, уметь измерять координаты рассеянного электрона с помощью прямого эксперимента. Но это заведомо не всегда возможно: интеграл в (41) берется по r от $-\infty$ до $+\infty$, для больших $|r|$ корень в экспоненте сделается мнимым, и соответствующие волны быстро затухают с ростом x. Как ловить такой электрон, как его измерять, мы не знаем. Поэтому поставим вопрос совершенно конкретно: то, что мы до сих пор рассматривали, еще не микроскоп. Чтобы получить микроскоп, мы поместим на пути рассеянных электронов линзу или систему линз и посмотрим, что даст линза на плоскости пластинки, какие заключения можно будет сделать по даваемому линзой изображению.

Через линзу идет сумма плоских волн (41), и нам следует теперь учесть действие оптического аппарата, действие линзы наиболее удобным способом. Это рассмотрение нужно произвести раз навсегда, чтобы в дальнейшем к нему уже не возвращаться. Поэтому мы предпримем сейчас небольшое отступление в область оптики.

Если поставить вопрос о том, какая оптическая система может отобразить плоскость на плоскость, то оказывается, что это будет *телескопическая система*[1]. Ее свойство в том и состоит, что падающая на нее плоская волна $e^{i(ry+\sqrt{k^2-r^2}x)}$ ($x = 0$ — плоскость предмета) преобразуется в плоскую же волну $e^{i\left[\frac{r}{s}y+\sqrt{k^2-\left(\frac{r}{s}\right)^2}x'\right]}$, где x' отсчитывается вдоль прежней оси x, но от плоскости

[1] [См. том II, статья 56.]

изображения, принимаемой за плоскость $x' = 0$. Это свойство математически выражают, говоря, что телескопическая система действует, как линейный оператор.

Может ли такая система давать изображение? Покажем, что может. Пусть на плоскости $x = 0$ имеем некоторую структуру, которую можно описать формулой

$$\int_{-\infty}^{+\infty} f(r)\, e^{iry} dr.$$

Тогда поле, падающее на линзу, имеет вид

$$\int_{-\infty}^{+\infty} f(r)\, e^{i\left(ry + \sqrt{k^2 - r^2}\, x\right)} dr. \tag{45}$$

Для $|r| > k$ волны затухают. Реальные линзы не пропускают таких волн. Следовательно, поле, прошедшее линзу, будет

$$\int_{-k}^{+k} f(r)\, e^{i\left[\frac{r}{s} y + \sqrt{k^2 - \left(\frac{r}{s}\right)^2}\, x'\right]} dr, \tag{46}$$

т. е. интеграл будет распространяться уже на промежуток в конечных пределах, а в плоскости изображения $x' = 0$ мы получим структуру

$$\int_{-k}^{+k} f(r)\, e^{i \frac{r}{s} y} dr. \tag{47}$$

Если бы и здесь был интеграл от $-\infty$ до $+\infty$, то изображение было бы в точности подобно предмету, но в другом масштабе ($1/s$ в показателе), т. е. мы получили бы идеальное изображение. Если k достаточно велико, настолько, что $f(r) \approx 0$ там, где r приближается к k, то изображение будет очень хорошим, почти идеальным. Большие k — это малые длины волн в оптике, большие импульсы у электронов в нашем случае.

Возвратимся теперь к измерению координаты частицы с помощью электронного микроскопа.

Итак, мы должны написать выражение для Ψ, преобразованное нашей линзой. Мы предположим для упрощения, что масса частицы велика, так что $\varepsilon^2 \ll 1$ и членом $\varepsilon^2 (p - r)^2$ под корнем можно пренебречь, так как он мал либо по сравнению с k^2, либо по сравнению с r^2. Тогда до линзы имеем

$$\Psi(x, y, z) = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} e^{\frac{i}{\hbar}\left[ry + (p - r) z + \sqrt{k^2 - r^2}\, x\right]} dr. \tag{48}$$

После линзы

$$\Psi(x', y, z) = \frac{1}{2\pi\hbar}\int_{-k}^{+k} e^{\frac{i}{\hbar}\left[\frac{r}{s}y+(p-r)z+\sqrt{k^2-\left(\frac{r}{s}\right)^2}x'\right]}dr. \qquad (49)^1$$

Заметим, что уже теперь видно, как перейти к приближенным измерениям. Это будет случай небольших k, не очень коротких волн. Мы наблюдаем электроны в плоскости $x' = 0$, т. е. имеем

$$[\Psi(x', y, z)]_{x'=0} = \frac{1}{2\pi\hbar}\int_{-k}^{+k} e^{\frac{i}{\hbar}\left[\frac{r}{s}y+(p-r)z\right]}dr = \frac{e^{\frac{i}{\hbar}pz}}{2\pi\hbar}\int_{-k}^{+k} e^{\frac{i}{\hbar}\left(\frac{y}{s}-z\right)r}dr. \qquad (50$$

Итак, взяв в качестве исходной волновой функции частицы собственную функцию импульса $\psi_0(z) = e^{\frac{i}{\hbar}pz}$, мы получили смесь — сумму произведений собственных функций импульсов частицы и электронов. Следовательно, измеряя импульс рассеянного электрона в $x' = 0$, мы не узнаем состояния рассеивающей диафрагмы-частицы. Наше устройство *не* является измерителем импульса.

Но является ли оно измерителем координаты? Чтобы проверить это, мы должны взять в качестве волновой функции частицы до измерения, $\psi_0(z)$, собственную функцию координаты

$$\psi_0(z) = \delta(z - z_1), \qquad (51)$$

и посмотреть, какая функция $[\Psi(x', y, z)|_{x'=0}$ получится в этом случае.

Вместо (38) мы имеем теперь

$$\text{при } x = 0 \qquad \Psi = \delta(z - z_1)\,\delta(z - y), \qquad (52)$$

или, если воспользоваться разложением (39),

$$\text{при } x = 0 \qquad \Psi = \frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} e^{-\frac{i}{\hbar}p(z_1-z)}dp\cdot\delta(z - y). \qquad (53)$$

Так как Ψ представляет собой интеграл, т. е. сумму отдельных членов вида (38), и так как мы знаем, какое $\Psi(x', y, z)$ получается для каждого такого члена [а именно, формула (49)], то мы можем

---

[1] Здесь, в отличие от (41), уже нет сохранения импульса [даже если мы не будем отбрасывать член $\varepsilon^2(p - r)^2$]. Объясняется это, конечно, тем, что некоторый импульс приняла на себя линза.

сразу же написать окончательный результат

$$\Psi(x', y, z) = \frac{1}{2\pi\hbar}\int\limits_{-\infty}^{+\infty} dp e^{-i\frac{p}{\hbar}z_1} \frac{1}{2\pi\hbar}\int\limits_{-k}^{+k} dr e^{\frac{i}{\hbar}\left[\frac{r}{s}y+(p-r)z+\sqrt{k^2-\left(\frac{r}{s}\right)^2}x'\right]}. \tag{54}$$

На плоскости $x' = 0$ это дает

$$[\Psi(x', y, z)]_{x'=0} = \frac{1}{2\pi\hbar}\int\limits_{-\infty}^{+\infty} dp e^{-i\frac{p}{\hbar}z_1} \frac{1}{2\pi\hbar}\int\limits_{-k}^{+k} dr e^{\frac{i}{\hbar}\left[\frac{r}{s}y+(p-r)z\right]}. \tag{55}$$

Если можно считать, что $k = \infty$, то

$$[\Psi(x', y, z)]_{x'=0} = \delta(z-z_1)\,\delta\left(\frac{y}{s}-z\right) = \delta(z-z_1)\,\delta\left(\frac{y}{s}-z_1\right). \tag{56}$$

Таким образом, если до измерения интересующая нас система обладала определенной координатой, $\psi_0(z) = \delta(z-z_1)$, то в конце концов мы получаем произведение функций от y' и z_1', причем $\varphi(y) = \delta\left(\frac{y}{s}-z_1\right)$ — собственная функция координаты. Измерять координату электрона (координату пятна на пластинке), т. е. y, мы умеем классически. Измерив же y, мы получаем, что координата частицы до соударения была $z_1 = \frac{y}{s}$. Значит, наш микроскоп, по крайней мере при $k = \infty$, действительно является измерителем координаты. Заметим, что после измерения мы получили $u(z) = \delta(z-z_1) = \psi_0(z)$, т. е. волновая функция частицы (существующая после измерения в пределах подсовокупности $y \sim \text{const}$) совпадает с исходной волновой функцией. Таким образом, измерения координаты повторимы.

ЧЕТВЕРТАЯ ЛЕКЦИЯ

(8.V 1939 г.)

Мы разобрали некоторые принципиальные вопросы квантово-механической теории измерений и иллюстрировали их двумя типичными примерами. Следует подчеркнуть, что мы разобрали не все вопросы, касающиеся измерений, а только некоторые, наиболее характерные. Так, например, мы брали ψ только в виде функции координат и совсем не рассматривали зависимости от времени, хотя уже понятия «до» и «после» измерения — временны́е. Взаимо-

действие считалось совершающимся мгновенно, хотя в действительности измерение требует времени и учет этого обстоятельства существенно усложняет рассмотрение. Вопрос о времени важен еще и потому, что в волновой механике он тесно связан с вопросом об энергии.

Вопрос этот трудный, и в этом семестре мы его не будем рассматривать, так как есть такие случаи, в которых время не играет существенной роли (например, диафрагма большой массы и малой скорости).

Мы имеем теперь уже достаточный материал для того, чтобы занять определенную позицию в полемике об основах волновой механики, происходившей между Эйнштейном, Бором и другими. Эта полемика относится еще к точным измерениям, и поэтому уместно остановиться на ней именно сейчас. Рассмотрим сначала один вопрос, я сказал бы, терминологического характера.

Говоря об измерении, мы хотели узнать, какое значение имела величина λ (оператор R с собственными функциями $\psi_i(x)$) в системе *I до ее взаимодействия* с системой *II*, в которой после взаимодействия мы умеем прямо измерять величину μ (оператор S с собственными функциями $\varphi_i(y)$). Мы назвали устройство измерителем λ, если, имея до опыта $\psi_0 = \psi_i(x)$, мы после опыта получаем $\Psi = u_i(x)\varphi_i(y)$. При этом функция $u_i(x)$ вовсе не обязательно должна быть собственной функцией R.

Но возможна и другая постановка вопроса. Имеем для системы *I* функцию $\psi_0(x)$, вводим взаимодействие с системой *II*, и пусть случается так, что

$$\Psi(x, y) = \sum_i a_i \psi_i(x)\varphi_i(y), \tag{57}$$

где ψ_i — собственные функции R, а φ_i — собственные функции S. В целом Ψ не является собственной функцией ни R, ни S, но для подсовокупности с фиксированным μ_k Ψ будет собственной функцией R, а именно, $\psi_k(x)$, т. е. значение λ будет λ_k. Таким образом, здесь другое понятие измерения: зная $\varphi_k(y)$, причем μ мы измеряем после взаимодействия, когда система *II* уже свободна и на систему *I* не влияет, мы можем *предсказать* результат измерения последствий взаимодействия для системы *I* (т. е. результаты измерений, произведенных над системой *I* после ее взаимодействия с системой *II*).

Я видоизменю теперь рассуждение Эйнштейна с тем, чтобы, не меняя его по существу, приблизить его к нашей схеме.

Если состояние системы описывается функцией $\psi(x)$, то, как учит волновая механика, вы не можете предсказать достоверно одновременные значения двух величин, которым соответствуют два некоммутирующих оператора. Например, измеряя координату системы, обладающей точным импульсом и соответственно описывае-

мой функцией $\psi(x) = e^{\frac{i}{\hbar}px}$, мы получаем новую $\psi(x)$, такую, что при определенном x мы утрачиваем p. Здесь, говорит Эйнштейн, нет возражений и даже более: *реальным* является либо p, либо x. Но если мне удастся указать метод измерения, позволяющий однозначно предсказать, не трогая систему, одновременно и p, и x, если это удастся сделать, то это будет противоречить волновой механике. С этим утверждением я согласен: указав такой способ, мы действительно имели бы противоречие с волновой механикой.

По сути дела, Эйнштейн ставит вопрос еще глубже (он прямо так не говорит, но я его толкую). Возможно, что описание при помощи волновой функции относится только к некоторому классу состояний и что существуют измерения, относящиеся к случаям, которые не охватываются таким описанием. Т. е. возможно, что теория с функцией ψ *не полна и ее можно дополнить так, чтобы не было принципа неопределенности*. Как именно следует дополнить теорию, чтобы не получить при этом противоречий, Эйнштейн не говорит, но он уверен, что дополнить ее можно.

Очевидно, Эйнштейн не знал, что в такой плоскости вопрос уже ставился. А именно, Нейман исследовал, возможно ли «суммарное, общее» описание при помощи ψ-функции так уточнить и пополнить «скрытыми» параметрами, чтобы не было соотношения неопределенности. Нейман доказал, что *нельзя* дополнить волновую механику таким образом, чтобы изжить принцип неопределенности. Конечно, сам Нейман так не говорит, он хорошо знает, что такое физическая теория: сегодня она есть, завтра ее нет. Он говорит, что нельзя отказаться от принципа неопределенности, не отказываясь от основ волновой механики. Когда-нибудь это, может быть, и придется сделать, но существенно, что нельзя выкинуть одно без другого.

Эйнштейн ставил вопрос все же не так глубоко. Ему сразу возразил Бор, и потом Эйнштейн признал свою ошибку.

Эйнштейн говорит следующее: возьмем сравнительно общий случай, когда после взаимодействия получается

$$\Psi(x, y) = \sum_i a_i \psi_i(x) \varphi_i(y).$$

Делая измерение над системой II, уже не взаимодействующей с I, я получаю собственное значение μ оператора S. Это дает мне *точное* значение λ_i для оператора R в системе I. Но эта же функция $\Psi(x, y)$ может быть разложена иначе:

$$\Psi(x, y) = \sum_i b_i \chi_i(x) \theta_i(y), \tag{58}$$

где $\chi_i(x)$ и $\theta_i(y)$ — собственные функции двух других операторов,

скажем R' и S'. Может случиться, что ψ_i и χ_i — собственные функции двух некоммутирующих операторов

$$RR' \neq R'R. \qquad (59)$$

Значит, делая измерения над системой II, один раз для S, другой раз для S', я могу предсказать *точные* значения для двух некоммутирующих операторов в системе I, которая моими измерениями прямо не затрагивается. В этом Эйнштейн и усматривает противоречие с основными посылками волновой теории и указание на ее неполноту.

То, что случай (57), (58), (59) возможен, Эйнштейн показывает просто на примере, и его пример есть не что иное, как уже рассмотренный нами микроскоп. Если до измерения частица-дырочка имела определенный импульс p, то на пластинке ($x' = 0$) мы получаем, согласно (50)

$$[\Psi(x', y, z)]_{x'=0} = \frac{1}{2\pi\hbar}\int\limits_{-\infty}^{+\infty} e^{\frac{i}{\hbar}[ry+(p-r)z]}ds. \qquad (60)$$

В отличие от (50), здесь положено $s = 1$ (нет увеличения) и $k = \infty$ (т. е. аппарат идеальный, измерение точное). Эйнштейн берет частный случай $p = 0$ и, кроме того, отсчитывает z и y не от одной точки, так что z заменяется на $z - z_0$. Последнего мы делать не будем, но p положим равным нулю. Тогда

$$\Psi(y, z) = \frac{1}{2\pi\hbar}\int\limits_{-\infty}^{+\infty} e^{\frac{i}{\hbar}(ry-rz)}dr. \qquad (61)$$

Итак, (61) действительно имеет вид (57): каждый множитель — собственная функция оператора *импульса*. Значит, если мы нашли импульс r для системы II, то для системы I импульс будет равен $-r$.

Но, согласно (39),

$$\Psi(y, z) = \delta(y - z) = \int\limits_{-\infty}^{+\infty} \delta(z - y')\,\delta(y' - y)\,dy', \qquad (62)$$

т. е. та же самая функция Ψ (y, z) может быть представлена как сумма произведений собственных функций оператора координаты. Значит, если мы нашли для системы II координату y', то она будет y' и у системы I. Тем самым доказательство Эйнштейна завершено.

Но в действительности это неверно. В чем же ошибка? Я увидел ошибку очень скоро и потом уже утратил способность ее не видеть.

Теперь мне даже трудно излагать дело так, будто ошибки нет. Сводится же она к следующему.

Электроны, попадающие в определенные точки плоскости изображения, не обладают определенными импульсами, и наоборот, положение электрона, обладающего определенным импульсом, не определено. Пока я ничего не измеряю, я имею Ψ (y, z), где представлены все z и все p. Произведя измерения, я могу из всей совокупности измерений выделить подсовокупность, в которой измерение некоторой определенной величины, характеризующей систему *II*, привело к некоторому определенному результату и в которой система *I* описывается волновой функцией. В этой подсовокупности, если импульс точно измерен, то координата — любая, и наоборот. Суть дела в том, что, выполняя измерения различных величин, относящихся к системе *II*, мы выделяем разные подсовокупности: мы либо фиксируем φ_i (y), т. е. импульс системы *II*, либо фиксируем θ_i (y), т. е. координату системы *II*[1]. В каждой из этих подсовокупностей для системы *I* имеет силу принцип неопределенности. Но одновременные точные значения импульса и координаты в *разных* подсовокупностях вполне допустимы, и никакого противоречия с волновой механикой здесь нет.

Рассматриваемое Эйнштейном состояние систем *I* и *II* описывается волновой функцией Ψ (y, z), но по отношению к каждой из систем в отдельности оно является смесью. Уточнить это описание можно, выделяя подсовокупность, описываемую волновой функцией ψ (x). Эйнштейн ошибся в том, что, говоря «я измеряю», он не указывает, к какой подсовокупности это относится. Таким образом, его упрек направлен не по адресу, а если направить его по адресу, то это не упрек.

Физически неправильно, когда Эйнштейн говорит: «Мы измеряем систему *II*, не затрагивая систему *I*». Спрашивается, откуда система *II* получила свой импульс? От столкновения с системой *I*. Значит, если мы берем только те случаи, когда система *II* обладает некоторым определенным импульсом, то мы берем тем самым лишь определенные удары со стороны системы *I*. Если же у системы *II* определенная координата, то она получила от системы *I* другие удары, или, точнее, удары, полученные ею от системы *I*, не являются определенными. Таким образом, здесь просто неправильно применена теория вероятностей, и никакого повода к пересмотру волновой механики возражение Эйнштейна не дает.

Перейдем теперь к другому вопросу, который в свое время сильно занимал теоретиков, но теперь уже отошел на второй план.

Представим себе полупространство, в котором есть тормозящее поле, или потенциальный барьер, и пусть на него падает по-

[1] [В этом случае в формулах (57) и (58) под координатой x системы *I* нужно понимать z в обозначениях разобранного примера.]

ток электронов. Можно найти Ψ во всем пространстве, и вот оказывается, что, как бы ни был высок барьер, $\Psi \neq 0$ и внутри, и позади него. Значит, электроны могут проникнуть внутрь барьера, где потенциальная энергия может сколь угодно превысить принесенную электронами кинетическую энергию. Это тот же вопрос, которого я уже коснулся ранее применительно к осциллятору. В том, что полная энергия меньше потенциальной, видели парадокс. Я его не вижу. В классике, где $E = V + T$, это, конечно, было бы парадоксом. Но здесь, пока мы не придали всему утверждению физического смысла, пока не сопоставили ему макроскопического измерения (скажем, пятно на пластинке, получаемое посредством «освещения» частицы), пока мы не имеем опыта, вообще нельзя сказать, парадокс это или нет. Во всяком случае, это очень интересное и большое отличие от классики, и наша задача будет теперь состоять в том, чтобы, взяв конкретный аппарат, например микроскоп, посмотреть детально, что физически означает случай $V > E$. Излагаемый ниже анализ этого вопроса провел М. А. Леонтович.

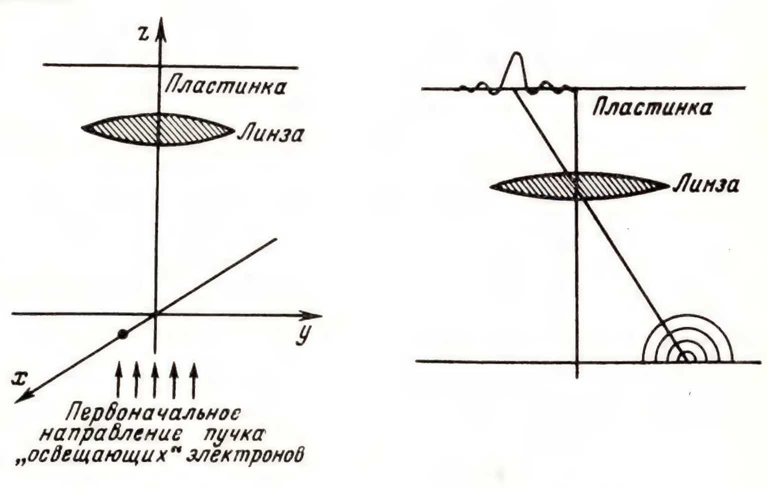


Рис. 6 **Рис. 7**

Для того чтобы рассмотреть случай частицы в силовом поле с большой потенциальной энергией полно и вместе с тем просто, возьмем сильно идеализированное, окарикатуренное устройство. Пусть частица массы M может двигаться только по оси x, а «освещающие» ее электроны — только в плоскости yz (рис. 6). Наша задача состоит в том, чтобы сравнить даваемое электронами изображение частицы при наличии и в отсутствие силового поля, в кото-

ром потенциальная энергия частицы V очень велика. До взаимодействия мы имеем

$$\Psi_0(x, y, z) = e^{i(k_x x + k_z z)}, \tag{63}$$

где k_x — импульс частицы, k_z — импульс освещающих электронов.

Пусть массы частицы и электрона одинаковы ($M = m$). Нам надо решить волновое уравнение

$$\frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2} + \frac{\partial^2\Psi}{\partial z^2} + \frac{2m}{\hbar^2}(E - V)\Psi = 0, \tag{64}$$

где E — полная энергия совокупности обеих систем (частицы и электрона), а $V \equiv V(x)$ — потенциальная энергия частицы в заданном поле. Это уравнение можно записать так:

$$\nabla^2\Psi + k^2\Psi = 0, \quad k^2 = \frac{2m}{\hbar^2}(E - V). \tag{65}$$

Мы предположим, что в достаточно большой области значений x $V(x)$ постоянно. Строго говоря, мы должны были бы писать $E - V - U$, понимая под U энергию взаимодействия электронов с частицей. Но мы предполагаем, что это взаимодействие проявляется лишь на очень малом расстоянии a между электронами и частицей. Расстояние R между ними, поскольку у электрона $x = 0$, а у частицы $y = z = 0$, есть

$$R = \sqrt{x^2 + y^2 + z^2}.$$

Примем, что при $R = a$ взаимодействие (отталкивание) электрона и частицы бесконечно сильное; это, как известно, сводится к тому, что $\Psi = 0$ при $R = a$. Итак, мы должны решить уравнение (65) при начальной падающей волне (63) и при условии

$$\text{при } R = a \quad \Psi = 0. \tag{66}$$

Это, очевидно, не что иное, как оптическая задача о дифракции плоской волны, падающей на идеально проводящий шар. Если шар очень мал, то решение (рассеянная волна) есть

$$\Psi = A\frac{e^{ikR}}{R}. \tag{67}$$

Такова функция Ψ *до линзы* электронного микроскопа. Мы уже теперь можем сказать, что будет при наличии и в отсутствие V. Если $E > V$ ($k^2 > 0$), то решение отлично от нуля даже на бесконечно больших расстояниях. Если же $V > E$ ($k^2 < 0$), то мы получаем затухающую волну $\frac{e^{-|k|R}}{R}$, которая сойдет на нет еще

до линзы, и изображения получить не удастся. Однако доведем задачу до конца.

Наш микроскоп двухмерный, и это несколько меняет дело по сравнению с пространственным случаем, где после линзы каждое пятно получается от элементарной сферической волны, падающей на линзу. В нашем случае рассеянные волны цилиндрические и элементарным решением является функция Ганкеля $H_0^{(1)}(\varkappa r)$, имеющая логарифмическую особенность в нуле и спадающая, как $1/\sqrt{r}$. при больших $\varkappa r$ (r — цилиндрический радиус-вектор). Это значит, что если до линзы мы имеем поле $H_0^{(1)}(\varkappa r)$, то на пластинке получим «пятно», т. е. поле вида (рис. 7)

$$\Delta(\varkappa y) = \frac{\sin \varkappa Ly}{\varkappa Ly}, \tag{68}$$

где параметр L характеризует нечто вроде ширины линзы, выраженной в длинах волн. Мы получили в пространстве до линзы поле (67). Если разложить это поле по $H_0^{(1)}$, то мы тотчас же сможем указать, какое оно дает изображение на пластинке.

Именно, до линзы рассеянная волна

$$\Psi = A\frac{e^{ikR}}{R} = iA\int\limits_0^\infty H_0^{(1)}(r\sqrt{k^2 - s^2})\cos sxds, \tag{69}$$

причем

$$r = \sqrt{y^2 + z^2}. \tag{70}$$

На пластинке каждая волна $H_0^{(1)}$ дает свое Δ, причем при бо́льших L Δ выглядит почти как δ-функция. В рассматриваемом сейчас вопросе, по существу, можно отвлечься от конечности размеров линзы и считать ее бесконечно большой. Тогда $L \to \infty$, Δ имеет вид δ-функции, т. е. каждая линия, служащая осью цилиндрической волны $H_0^{(1)}(\varkappa r)$, даст на пластинке также линию. От этой идеализации нам все равно придется отказаться (см. лекцию 5), когда мы будем рассматривать приближенные измерения. Поэтому мы не будем предполагать, что $L \to \infty$, а будем пользоваться формулой (68).

Изображение на пластинке дают лишь незатухающие волны, т. е. только те волны, у которых $\varkappa = \sqrt{k^2 - s^2}$ действительно. Для мнимых $\varkappa$ изображения нет, Δ экспоненциально стремится к нулю. Таким образом, поле (69) дает на пластинке поле (при произвольном L)

$$\Psi = iA\int \Delta(y\sqrt{k^2 - s^2})\cos sx\,ds, \tag{71}$$

где интеграл распространен на действительные значения $\sqrt{k^2 - s^2}$, т. е. при действительном k от $s = 0$, до $s = k$.

Мы видим, что если $V = 0$, то при всякой энергии $E = \frac{\hbar^2}{2m}(k_x^2 + k_z^2)$ k действительно и в интеграле (71) всегда найдутся члены, для которых корень $\sqrt{k^2 - s^2}$ действителен. Следовательно, в случае $V = 0$ изображение получится всегда.

Если же потенциальная энергия настолько велика, что $k^2 < 0$, то на пластинке

$$\Psi = iA \int \Delta (iy \sqrt{|k|^2 + s^2}) \cos sx \, ds, \qquad (72)$$

т. е. ни при каких s аргумент функции Δ не становится действительным и изображения не получится. Можно ли вообще увидеть частицу с помощью нашего микроскопа, если $V \neq 0$? Очевидно, это становится возможным всякий раз, когда энергия освещающих электронов взята настолько большой, что $E > V (k^2 > 0)$.

Как видите, в результате произведенного анализа оказалось, что никакого парадокса нет. Действительно, бывает, что $\Psi \neq 0$ там, где $V > E$. Но что это значит? Из разобранного примера мы видим, что утверждение о местонахождении частицы в области больших V имеет смысл, если можно осветить частицу электронами со столь большой энергией, что $E > V$. Но тогда нет парадокса с энергией.

ПЯТАЯ ЛЕКЦИЯ

(20.V 1939 г.)

Наши общие положения заключались в том, что всякая физическая теория состоит из двух (связанных друг с другом) частей: математической части, оперирующей символами (числами, операторами и т. п.), и измерений — рецептов, связывающих эти символы с объектами природы. Только обе эти стороны, взятые вместе, позволяют давать ответы на физические вопросы.

Под этим углом зрения мы рассмотрели некоторые общие вопросы измерений, а затем перешли к отдельным конкретным задачам. Располагая этим материалом, мы можем подойти теперь и к другим проблемам. В их числе имеется такой фундаментальный вопрос, как толкование соотношения неопределенностей. Многие склонны считать это соотношение тем центральным пунктом, который отличает волновую механику от классики.

Однако мне представляется нецелесообразным начать теперь дискуссию этого вопроса во всей полноте с тем, чтобы потом где-то ее прервать. Быть может, лучше будет, в порядке подготовки к следующему семестру, коснуться лишь некоторых сторон вопроса, и то не в общем случае, а показать на примере электронного микроскопа, в чем здесь дело и какие здесь встречаются неточные утверждения.

Сначала один общий вопрос — о происхождении соотношения неопределенностей. Мы видели, что самое лучшее и самое большее, что может сделать волновая механика,— это дать, во-первых, функцию ψ, во-вторых, рецепты того, как следует сопрягать получаемые из ψ числовые величины с результатами измерений различных физических величин. Импульсу при этом приписывается оператор $\frac{\hbar}{i}\frac{\partial}{\partial x}$, и собственные значения этого оператора и являются возможными значениями импульса. *Уже здесь, т. е. в определениях, содержится соотношение неопределенностей. По определению*, неточность в координате Δx берется из естественной меры разброса

$$(\Delta x)^2 = \int\limits_{-\infty}^{+\infty} (x - \bar{x})^2 \, |\psi(x)|^2 \, dx. \tag{73}$$

Точно так же, по *определению*,

$$(\Delta p)^2 = \int\limits_{-\infty}^{+\infty} (p - \bar{p})^2 \, |g(p)|^2 \, dp, \tag{74}$$

где $g(p)$ — коэффициенты разложения ψ по собственным функциям оператора импульса. Отсюда чисто математически вытекает, что [1]

$$\Delta x \cdot \Delta p \geqslant \frac{\hbar}{2}.$$

Таким образом, можно сказать, что соотношение неопределенностей заключено в математическом аппарате, в математической символике квантовой механики.

С другой стороны, говорят (Дирак ставит это во главу угла), что соотношение неопределенностей обусловливается *возмущением измеряемой системы*, неизбежно испытываемым ею в акте измерения; всякое измерение x дает разброс p_1 и обратно.

Грубо говоря, имеются эти две точки зрения на происхождение соотношения неопределенностей. Ясно, однако, что это не противо-

---

[1] [См., например, *В. Гейзенберг*. Физические основы квантовой механики; *Д. И. Блохинцев*. Введение в квантовую механику, § 15].

поставление, а две стороны одного и того же. Полная теория необходимо содержит и математический аппарат, и рецептуру измерений. Дирак подчеркивает второе, я здесь — первое.

Могло бы случиться, что в математическом аппарате заложено соотношение неопределенностей, в то время как измерения одновременно давали бы точные значения и x, и p. Тогда в теории было бы заключено внутреннее противоречие. Но дело спасают возмущения при измерении: x и p можно было бы одновременно измерять точно, если бы измерительные устройства были таковы, что функция ψ не претерпевала бы изменений. Но этого нет, ибо акт измерения *согласно* теории (или, лучше сказать, теория *так* и была построена) неизбежно изменяет ψ. Значит, указанные две точки зрения не противоречивы, а что считать за исходное, это до известной степени дело вкуса. Можно сказать, что теория сразу же строится так, чтобы она отражала экспериментальные факты. Но можно сказать и так, что я постулирую теорию, а затем смотрю, подходит она или не подходит к природе.

Во всяком случае, вы видите, что уравнение Шрёдингера здесь не при чем, что соотношение неопределенностей возникает еще до уравнения Шрёдингера.

Итак, если я хочу точно измерить x, то, вообще говоря, я должен при этом *совершенно* разрушить $\psi_0(x)$ [1]. Именно так обстоит дело при точных измерениях, о которых мы до сих пор только и говорили. Но нельзя ли пойти на компромисс: ценой не вполне точного знания x исказить $\psi_0(x)$ не до конца, т. е. так, чтобы какие-то основные черты исходного состояния сохранились? Другими словами, как подойти к вопросу об измерениях с допуском, к вопросу о *приближенных* измерениях, стоявшему очень остро и в волновой механике?

В классике приближенное знание величины, как правило,— минус. Всегда желательно знать все величины возможно точнее, и в классике это не встречает принципиальных препятствий. В волновой же механике к этому имеется принципиальное препятствие. Приближенные измерения мы теперь и рассмотрим. Мы ограничимся при этом примером микроскопа — прибора, позволяющего в идеальном случае точно измерять координату.

Разумеется, вопрос об измерениях шире и сложнее, чем то, что мы рассматривали. К счастью, с микроскопом все обстоит хорошо, хотя если начать разговор и об измерении импульса, то появляются некоторые тонкости. Ландау и Пайерлс обратили внимание на то, что представление в x-пространстве совершенно равнозначно представлению в p-пространстве. Символически переход между этими

[1] [Если только $\psi_0(x)$ не является δ-функцией от x. См. конец 3-й лекции, стр. 360.]

представлениями можно записать так:

$$\left(x, \frac{\partial}{\partial x}\right) \to \left(-\frac{\partial}{\partial p}, p\right). \quad (75)$$

Можно спросить, откуда же берется несимметрия x и p в некоторых задачах? Ответ заключается в том, что взаимодействие обычно является функцией от x, а не от p, т. е. природа устроена так, что не всякие гамильтонианы H существуют. Обычно мы берем H из грубой классической модели и переносим его в микромир. Откуда же мы знаем, что в микромире не может быть других H? Заранее это ниоткуда не следует, но принято считать, что других гамильтонианов в микромире не существует.

Возвращаясь к вопросу о приближенных измерениях, необходимо прежде всего отметить, что в этом случае нужно уже иначе строить теорию. При точных измерениях сама функция $\psi(x)$ имела смысл только потому, что существуют измерительные устройства, дающие точное значение x. Теперь речь идет о неточных значениях x. Как же следует изменить микроскоп, чтобы сделать его приближенным измерителем, и чего мы можем достичь таким путем?

Прежде чем заняться этим вопросом, я хотел бы сначала дать критику того, что обычно говорят по поводу микроскопа и что всегда меня шокировало.

Вот обычное рассуждение: если апертура равна φ (рис. 8), то сколь угодно малая диафрагма дает в плоскости изображения кружок, средний размер которого $\lambda/\sin\varphi$. Значит, я не могу определить положение диафрагмы с меньшей ошибкой, чем $\Delta x \sim \lambda/\sin\varphi$. Далее, так как существует комптон-эффект (сохранение импульса при взаимодействии фотона и диафрагмы), то диафрагма может заимствовать от фотона любой импульс в пределах от 0 до $\frac{h\nu}{c}\sin\varphi$, т. е. импульс диафрагмы известен с точностью до $\Delta p \sim \frac{h\nu}{c}\sin\varphi$. Следовательно,

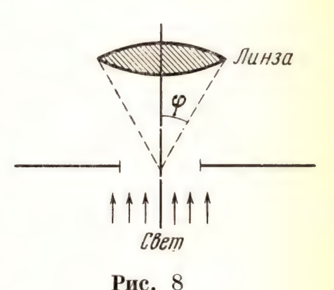


Рис. 8

$$\Delta x \cdot \Delta p \sim h.$$

Не знаю, может быть, с моей стороны здесь проявляется известный педантизм, но каждое слово в этом рассуждении, по-моему, ошибочно (хотя по существу все так и есть).

Во-первых, ссылаются на то, что разрешающая сила не может быть больше $\lambda/\sin\varphi$, т. е. что мы не можем различить *два* более

близких кружка. Но *одна* точечная диафрагма дает *один* кружок. Что же мне мешает точно определить координату его центра? Очевидно, с классической оптикой здесь ничего нельзя получить, надо сразу же говорить об *отдельном фотоне*, а не о *кружке* почернения фотопластинки, получающемся под действием многих фотонов. Место, куда попадает отдельный фотон, действительно не определено однозначно.

Во-вторых, как из того факта, что фотон попал в данную точку, заключают о координате диафрагмы? Почему точку на пластинке, в которую попал фотон, проектируют через центр линзы? Давайте лучше не думать об этом.

Наконец, получают $\Delta x \cdot \Delta p \sim h$. Но существует общая теорема, согласно которой ширина спектра разрывной функции бесконечна [1]. Между тем берут диафрагму с резкими краями (дырочку в экране), а тогда, как легко убедиться и прямым расчетом, $\Delta p = \infty$. Таким образом, в таком неудачном примере мы имеем попросту

$$\Delta x \cdot \Delta p = \infty.$$

Мы оставим теперь в стороне все такие и подобные соображения и, пользуясь оптической математикой (но не физической оптикой), рассмотрим микроскоп с конечным значением отношения $\lambda/\sin\varphi$ строго. Именно этим разбираемая нами сейчас задача отличается от случая $\lambda/\sin\varphi \to 0$, справедливого, когда микроскоп дает точное знание координаты. В остальном все расчеты прежние. Нас интересуют два вопроса: 1) что дает нам такой микроскоп на пластинке и 2) в каком состоянии находится наблюдаемая частица после измерения.

Если мы повторим весь вывод, содержащийся в формулах (36) — (50), взяв для исходного состояния частицы-дырочки не $e^{\frac{i}{\hbar}pz}$, а некоторую функцию $\psi_0(z)$, то, как нетрудно видеть, получим на пластинке (полагая $s = 1$)

$$\Psi(y, z) = c\psi_0(z) \int e^{\frac{i}{\hbar} r(y-z)} dr, \tag{76}$$

где y — координата пятна, оставляемого электроном на пластинке. Можно ввести обозначение

$$\frac{r}{\hbar} = \frac{2\pi}{\lambda} \sin\varphi. \tag{77}$$

Здесь φ не что иное, как угол, образуемый направлением полета

---

[1] [См., например, *С. М. Рытов*. Труды Физического института им. П. Н. Лебедева Академии наук СССР, т. II, вып. 1, 1938, стр. 59 и 121.]

рассеянного электрона с осью микроскопа. Если мы интегрируем по r от $-\infty$ до $+\infty$, то получаем точный микроскоп

$$\Psi(y, z) = c'\psi_0(z)\,\delta(y - z). \tag{78}$$

Но теперь мы будем интегрировать по r от $-a$ до $+a$, так как линза или какая-то другая диафрагма ограничивает угол[1]. Тогда

$$\Psi(y, z) = c''\psi_0(z)\,\frac{\sin\frac{a}{\hbar}(y-z)}{\frac{a}{\hbar}(y-z)}. \tag{79}$$

Если теперь рассчитать Δp, то и получается $\Delta p = \infty$.

Разумнее исходить из другого микроскопа (принципиально легко осуществимого), у которого происходит плавное ослабление волн на «краях» диафрагмы, тем большее, чем больше r (или φ). Возьмем фактор ослабления в виде e^{-r^2/k^2}. Тогда мы получим вместо (76) и (79)

$$\Psi(y, z) = c\psi_0(z)\int\limits_{-\infty}^{+\infty} e^{-\frac{r^2}{k^2}}e^{\frac{i}{\hbar}r(y-z)}dr = c'''\psi_0(z)\,e^{-\frac{k^2}{4\hbar^2}(y-z)^2}. \tag{80}$$

В этом наиболее выгодном случае получается $\Delta x\cdot\Delta p = \frac{\hbar}{2}$.

В частности, если мы заведомо знаем, что частица находилась в z', т. е. $\psi_0(z) = \delta(z - z')$, то

$$\Psi(y, z) = c\delta(z - z')\,e^{-\frac{k^2}{4\hbar^2}(y-z)^2}. \tag{81}$$

Вероятность того, что электрон попал на пластинку в точку $y = y'$, будет тогда $|f(y', z')|^2$, где $f(y, z) = e^{-\frac{k^2}{4\hbar^2}(y-z)^2}$.

Поставим теперь вопрос иначе. Рассмотрим те случаи, в которых координата пятна, оставляемого электроном на пластинке, имела некоторое фиксированное значение y. Какую координату z имела при этом частица-щель? Однозначного ответа нет. Однозначный ответ получается только при $k \to \infty$, когда (80) переходит в

---

[1] Выражая с помощью (77) r через φ и полагая, что $\frac{\lambda a}{2\pi\hbar} \ll 1$, можно сказать, что в случае микроскопа с отверстием диафрагмы конечных размеров мы интегрируем по углу φ полета рассеянного электрона от $\varphi = -\frac{\lambda a}{2\pi\hbar}$ до $\varphi = +\frac{\lambda a}{2\pi\hbar}$.

(78), т. е. когда координата частицы z заведомо равна координате пятна y.

Мы можем, однако, определить, какова вероятность того, что $z = z'$, если известно, что $y = y'$. Итак, сначала мы предполагали известным $\psi_0(z)$; тогда вероятность события $z = z'$ есть $|\psi_0(z')|^2$. Вероятность же того, что при $z=z'$ получается $y=y'$, есть $|f(y', z')|^2$. Но сейчас наша цель состоит в том, чтобы получить хотя бы некоторое приближенное знание о координате z частицы по результатам измерений с помощью неточного микроскопа (т. е. микроскопа с конечной апертурой). Поэтому сейчас мы спрашиваем: какова вероятность того, что $z = z'$, если я обнаружил $y = y'$? Это типичная задача Байеса о вероятности a posteriori [1].

Имеются урны A_i, и вероятность попасть на урну A_i есть $W(A_i)$. В урнах лежат шары разных цветов, и вероятность того, что, попав в урну A_i, я вытащил белый шар, есть $W(a, A_i)$. Спрашивается, если я вынул белый шар, то какова вероятность того, что это была урна A_i? Обозначим искомую вероятность через $W(A_i, a)$. Тогда, как известно,

$$W(A_i, a) = W(A_i)\frac{W(a, A_i)}{\sum\limits_j W(A_j)\,W(a, A_j)}. \tag{82}$$

Замечу, что Мизес в своей книжке [2] резко восстает против такой терминологии (вероятности a posteriori или вероятности гипотез, причин), называя ее метафизической и указывая, что здесь мы имеем просто специальный выбор подсовокупности. Это верно, но результата (82) не меняет.

В нашем случае $W(A_i)$ — это $|\psi_0(z')|^2$, а $W(a, A_i)$ — это $|f(y', z')|^2$. Таким образом, вероятность нахождения частицы в $z = z'$, когда в изображении $y = y'$, есть

$$W(z') = \frac{|\psi_0(z')|^2 \cdot |f(y', z')|^2}{\int\limits_{-\infty}^{+\infty} |\psi_0(z')|^2 \cdot |f(y', z')|^2\,dz'}. \tag{83}$$

В предельном случае точного микроскопа, когда $k \to \infty$, $f(y, z) = \delta(y - z)$, и мы получаем из (83), что $W(z')$ пропорционально $\delta(y' - z')$, т. е. в точном случае мне не нужно знать $\psi_0(z)$. В приближенном же случае знать $\psi_0(z)$ нужно. Это и понятно: если одно только $W(a, A_5)$ отлично от нуля (а это как раз и соот-

---

[1] [См., например, *С. Н. Бернштейн*. Теория вероятностей. Изд. 4-е, ОГИЗ, 1946, стр. 76.]

[2] [По-видимому, имеется в виду *R. v. Mises*. Wahrscheinlichkeit, Statistik und Wahrheit. Springer, 1928, стр. 95.]

ветствует случаю $f = \delta(y - z)$, все же остальные $W(a, A_i)$ равны нулю, то, какой бы шар мы ни вынули, он будет из пятой урны. Или вообще если в A_1 только красные шары, в A_2 — только зеленые и т. д., то знание цвета однозначно определяет урну.

Что касается состояния частицы после измерения, то ведь с этого мы и начали: до измерения была $\psi_0(z)$, после измерения, т. е. при известном $y = y'$,

$$\Psi(y', z) = c\psi_0(z) f(y', z). \tag{84}$$

Если $f(y', z)$ — очень острая δ-образная функция (т. е. микроскоп хороший, k или a велики), то состояние коренным образом меняется. Если же $f(y', z)$ — расплывчатая функция от z, такая, что в пределах тех значений z, в которых сосредоточены отличные от нуля значения $\psi_0(z)$, $f(y', z)$ почти неизменна, то $\Psi(y', z)$ подобна $\psi_0(z)$, состояние измеряемой системы меняется мало. Конечно, лишь в таком специальном случае $\Psi(y', z)$ просто пропорциональна $\psi_0(z)$. Он имеет место благодаря коммутации энергии взаимодействия с z и y, но существа дела это не меняет. Сводится же оно к тому, что, выбирая аппарат, делая его более или менее точным, мы можем либо точно измерять z и тогда совершенно расстраиваем систему, либо же можем сохранить основные черты состояния системы и тогда не точно знаем z. Этим мы можем *управлять* по желанию, по своей надобности. Таким образом, приближенный характер результата измерений, в отличие от классики, здесь не является минусом.

Последнее замечание: что я сделал, чтобы перейти от точных измерений к приближенным? Я прикрыл диафрагму микроскопа, *не влияя* на акт соударения электрона с частицей. Как же получилась другая волновая функция? Ответ всегда надо искать в том, что́ же именно хотят сказать. Сплошь и рядом недоумение порождается просто неточными выражениями. Да, соударения мы не изменили, но мы *выбрали другую подсовокупность опытов*, мы выбрали только те из рассеянных электронов, которые летят внутри определенного угла; поэтому нет ничего удивительного в том, что и функция Ψ стала иная.

Мы разобрали некоторые принципиальные вопросы, относящиеся к квантовомеханическим измерениям, выяснили сущность таких измерений и потом на примерах видели, как эти принципы применяются к той или иной конкретной задаче.

Я хочу, однако, подчеркнуть в согласии с тем, что я сказал уже во вступлении, что этими рассуждениями далеко не исчерпываются все вопросы, относящиеся к измерениям.

Укажу, например, на следующее. Мы предполагали, что измерительное устройство таково, что если наша измеряющая система находится в собственном состоянии, описываемом функцией $\psi_i(x)$ с собственным значением λ_i, то после взаимодействия $\Psi(x, y) = u_i(x)\,\varphi_i(y)$, т. е. что есть однозначное соответствие между λ_i и μ_i. Я уже тогда упомянул, что это не необходимо, что достаточно однозначного соответствия, при котором одному λ соответствует несколько μ, но так, что каждому μ соответствует только одно λ.

Более существенный вопрос — это вопрос о времени. В нашей постановке вопроса время играло второстепенную роль. Другими словами, мы не учитывали, например, продолжительности измерений. Есть класс явлений, где такое рассмотрение законно. Но какую роль время играет при измерениях? Этот глубокий и интересный вопрос оказывается неизбежно связанным с вопросом об энергии, о принципе сохранения энергии. Соблюдается ли закон сохранения энергии в каждом отдельном акте взаимодействия? Шенкланд думал, что наблюдал нарушения этого закона [1]. Надо сказать, что в своей классической форме закон сохранения в волновой механике неприменим. Просто понятия в классике такие, которых здесь нет. Все же правильная теория предсказывает, что если пользоваться обычными формулировками, то в опытах Шенкланда должно быть сохранение энергии в каждом отдельном акте. Таким образом, *если бы* он был прав, то это было бы опровержением волновой механики. Но он оказался неправ.

Другой вопрос — о связи или переходе от волновой механики к классической. Обычно указывают на теорему Эренфеста, но я думаю, что это не единственный путь. Ведь в классике есть не только детерминистическая ньютонова механика, к которой ведет теорема Эренфеста, но и статистическая теория. Как связаны микро- и макростатистики — это особый вопрос.

В этом семестре на всех подобных вопросах останавливаться я не могу.

[1] [Имеются в виду впоследствии опровергнутые эксперименты, в которых якобы было доказано, что в отдельных актах взаимодействия фотонов с электронами законы сохранения энергии и импульса не соблюдаются, но что они верны в среднем для большого числа актов.]

ДОПОЛНЕНИЕ[1]

Ряд положений квантовой механики чрезвычайно сильно отличается от того типа положений, к которым мы привыкли в классике. Это отличие настолько велико, что часто говорят о новом физическом мировоззрении. Я хотел бы уже сейчас отметить (к этому я вернусь ниже), что, по моему мнению, *принципы* построения квантовой теории, или, если так можно выразиться, структура той рамы, которой квантовая теория обрамляется, та же, что и в любой другой физической теории. Но нельзя отрицать, что структура самой картины весьма отлична от классики, и утверждение, что мы здесь имеем дело с новым физическим мировоззрением, вряд ли можно считать преувеличением. Именно поэтому небезынтересно постараться, во-первых, проанализировать те моменты, которые определяют новизну квантовой концепции, и, во-вторых, уяснить себе связь между отдельными ее положениями. Такой анализ, мне кажется, не лишен интереса и сам по себе. Но, кроме того, он дает ответ на следующий вопрос, который всегда возникает и, естественно, должен возникать. Это вопрос относительно того, насколько тесно связаны те стороны теперешней, фактически построенной квантовой теории — я имею в виду в основном нерелятивистскую теорию,— которые обусловливают ее бесспорную силу и громадное значение для решения конкретных проблем, с теми ее общими положениями, которые придают ей ее революционный характер. Среди физиков встречается мнение, что эти последние лишь провизорны и что в будущем удастся, не нарушая действенности квантовой теории, свести ее общие положения к положениям классического типа. Ответ на поставленный выше вопрос должен внести здесь ясность.

Я хотел бы указать, что этот вопрос по существу был поставлен и разрешен Нейманом. Результаты его работы обычно приводят без доказательства и в чрезвычайно схематичной форме. Например, говорят, что Нейман *доказал* невозможность построения физической теории на детерминистической основе. По моему мнению, такая и ей подобные формулировки мало что говорят. Так как, насколько я могу судить, исследования Неймана мало известны, то я считал бы не бесполезным вернуться к этому и аналогичным вопросам. Базируясь в основном на работах Неймана, я, однако, в ряде доказательств, а также в выборе для рассмотрения тех или иных вопросов не всегда буду придерживаться его изложения.

Но предварительно я хотел бы сделать следующее замечание. Квантовая теория, во всяком случае в ее современном виде, по

---

[1] [См. примечание на стр. 325.]

существу теория атомистическая в том смысле, что она относится к явлениям в микромире, которые, согласно атомистической гипотезе, лежат в основе явлений макромира. Здесь не место останавливаться на вопросе о том, почему мы кладем в основу физического миросозерцания именно атомистическую теорию. Несомненно, что в настоящее время атомизм считается обоснованным больше, чем когда бы то ни было прежде, и по этому поводу никаких вопросов, вообще говоря, не возникает. Возражение вызывает лишь отсутствие в относящейся к микромиру квантовой теории того, что называется наглядностью. Это отсутствие наглядности есть одна из причин, заставляющих иногда считать всю квантовую концепцию, как это уже указывалось выше, провизорной. Однако, на мой взгляд, эта точка зрения весьма поверхностна и основана на недоразумении. Не «ненаглядность» сама по себе придает квантовой теории ту новизну, специфичность и то отличие от классики, о которых идет речь в дальнейшем.

По поводу отсутствия наглядности я хочу все же заметить следующее. Понятие наглядности вообще чрезвычайно условно. В него входят, по существу, два элемента. Во-первых, чтобы что-нибудь было наглядно, оно должно быть привычно; в этом и лежит условность этого понятия. Затем необходимо, чтобы наглядная концепция связывалась с вещами, могущими быть *непосредственно* чувственно воспринятыми.

Вообще говоря, нет никакого основания требовать от новой теории, чтобы она удовлетворяла этим двум условиям. Насколько такого рода требование не обоснованно, с одной стороны, и подвержено эволюции — с другой, хорошо видно на известном примере развития классических теорий. Здесь механические явления были прежде других изучены, сделались привычными. Эти же явления легко воспринимаются зрением, осязанием и т. д. И вот, начиная примерно со времен Ньютона, ко всякой теории, хотя бы она относилась уже явно к немеханическим явлениям, как-то: к оптической теории, теории тепла и т. д., — предъявлялось требование наглядности, выражавшееся в конечном счете в требовании сведе́ния оптических, тепловых и т. д. явлений к явлениям механическим. В своем знаменитом «Traité de la lumière» Гюйгенс, например, писал: «... au moins dans la vraie philosophie, dans laquelle on conçoit la cause de tous les effets naturels par des raisons de mécanique. Ce qu'il faut faire à mon avis, ou bien renoncer à toute esperance de ne jamais rien comprendre dans la physique». (Изд. Gautier-Villars, 1920, стр. 3) [1].

[1] [«... по крайней мере в истинной философии, в которой источник всех явлений природы усматривают в механических причинах. На мой взгляд, это и нужно делать или же отказаться от всякой надежды когда-либо понять что-нибудь в физике».]

Ньютон в предисловии к первому изданию «Principia» говорит (перевод А. Н. Крылова, стр. 3): «Было бы желательно вывести из начал механики и остальные явления природы, рассуждая подобным же образом, ибо многое заставляет меня предполагать, что все эти явления обусловливаются некоторыми силами, с которыми частицы тел вследствие причин, покуда неизвестных, или стремятся друг к другу и сцепляются в правильные фигуры, или взаимно отталкиваются и удаляются друг от друга. Так как эти силы не известны, то до сих пор попытки философов объяснить явления природы и остались бесплодными». То, что Ньютон имеет здесь в виду и световые явления, становится ясным в отделе XIV первой книги «Principia», где он подробно излагает с этой точки зрения отражение и преломление света, а в отделе VIII второй книги опять упоминает о свете в связи с теорией распространения волн.

И в XX в. лорд Кельвин говорил, что для него понять физическое явление — значит построить механическую его модель.

Если вдуматься в это требование, которое казалось естественным в течение ряда веков, то трудно найти для него какое бы то ни было обоснование. Действительно, теоретическая механика — по существу это механика Ньютона — это система некоторых определенных дифференциальных уравнений второго порядка между математическими величинами (x, y, z, t) в связи с рецептами для сопоставления этих величин с физическими объектами.

И то и другое, т. е. и указанные выше специальные дифференциальные уравнения, и специальные рецепты, приурочено к определенному классу явлений, именно к тем, которые мы называем механическими, т. е. к явлениям, связанным с перемещениями, скоростями и т. д. макроскопических тел.

Оптические, тепловые и другие явления заведомо в этот класс не входят. Весьма естественно, что и для этих явлений могут быть построены теории по такому же принципу, как и в механике, т. е. могут быть найдены математические соотношения между некоторыми величинами в связи с рецептами измерений соответственных величин. Но совершенно не очевидно, и более того — маловероятно, чтобы эти соотношения, например дифференциальные уравнения, регулирующие оптические явления, были тождественны с механическими уравнениями Ньютона. А ведь именно на это сводится требование «наглядности», т. е. механического объяснения оптических и других явлений [1].

И действительно, как известно, в классике от такого требования, которое является чистым предрассудком, отказались совер-

[1] Известное замечание, что в конце концов все наши измерения сводятся на констатацию пространственных и временных совпадений, к этим вопросам непосредственного отношения не имеет.

шенно. Пример — максвелловская теория электромагнитных и оптических явлений.

Я думаю, всякий физик теперь согласится с тем, что максвелловская теория автономна (она вполне удовлетворяет тем требованиям, которые мы разумным образом к теории можем предъявлять), мы к ней привыкли и никакой дальнейшей дополнительной наглядности, как-то: сведéния на механику, которое оказалось невозможным,— не требуем; в этом смысле мы считаем ее совершенно равноценной механике Ньютона.

Возвращаясь к квантовой механике, мы видим, что здесь повторяется mutatis mutandis, аналогичная ситуация. Прежняя, доквантовая, классическая атомистика молчаливо всегда предполагала (по-видимому, в угоду наглядности), что в микромире справедливы те же соотношения, какие воспринимаются нами в макромире. Те, кто жалуется на «ненаглядность» квантовой механики, полагают, по-видимому, справедливым требование, чтобы этому пути следовала атомистическая теория, и считают недостатком квантовой теории то, что она по этому пути не идет. Я думаю, это неправильно. Не недостатком, а достоинством квантовой механики следует считать то, что она порвала с предрассудком о тождественности закономерностей в микромире и закономерностей в макромире и тем самым и с предрассудком о желательности «наглядности» в описании явлений в микромире.

Всякая атомистическая теория, в том числе и квантовая, ставит себе в основном задачу объяснить наблюдаемые закономерности в макромире искомыми закономерностями в микромире. Почему такое сведéние нас удовлетворяет, почему мы его называем объяснением процессов макромира,— это другой вопрос. На нем я останавливаться здесь не могу. Скажу только, что, по моему мнению, для такой точки зрения веские основания привести можно. Но как бы там ни было, а из этой формулировки задачи мы должны исходить. Нельзя отрицать, что атомная теория может в известном смысле, по крайней мере частично, решать эту задачу и в том случае, если эта теория исходит из положения, что атомы и молекулы подчиняются тем же законам, как и макротела, т. е., например, что они упруги, что для них справедливы законы движения Ньютона и т. д. На этой основе строилась, например, классическая кинетическая теория газов. Правда, нужно оговориться, что кинетическая теория газов вводит существенные статистические *добавочные* предположения, которые чужды механике Ньютона (но и на этом вопросе я сейчас останавливаться не буду). Как бы там ни было, *предполагалось*, что для молекул справедлива механика Ньютона и что в основном она определяет явления.

Нельзя, конечно, отрицать, что классическая кинетическая теория дала много интересных результатов. Но, с другой стороны,

нужно признать, что всякая атомистическая теория, оперирующая в качестве молекул или атомов объектами, которым приписываются свойства и поведение макротел, не может быть признана удовлетворительной. Мы ведь ставим себе в этих теориях задачу — объяснить свойства и поведение макротел поведением и свойствами конституирующих их микротел. Но тогда непоследовательно этим последним приписывать свойства, которые мы намереваемся объяснить.

Если же — и это делает квантовая теория — признать, что микротела подчиняются другим законам, чем макротела, для объяснения свойств которых эти микротела вводятся, такой regressus[1] по существу ad infinitum[2] отпадает. Конечно, можно и здесь спросить себя, нельзя ли ввести в рассмотрение дальнейшие микро-микротела. Однако никакой необходимости в такой постановке вопроса здесь уже нет, в то время как при доквантовой постановке она неизбежно возникает.

Резюмируя, можно сказать следующее.

Уже в классике долго держался предрассудок, что задача теоретической физики заключается в сведе́нии всех явлений к механическим. Мы видели, что это действительно не что иное, как предрассудок, возникший благодаря тому, что механические явления были раньше других изучены и поэтому более привычны.

Как и следовало ожидать, механическая интерпретация оптических, электромагнитных и других явлений фактически оказалась невыполнимой. Мы привыкли, например, к максвелловской теории и уже никакой потребности в таком сведении не ощущаем.

По отношению к квантовой теории ситуация аналогична. Выставляется требование свести законы микромира на привычные классические законы, относящиеся к макромиру. Это требование, переходящее в конечном счете в требование, чтобы здесь были справедливы те же дифференциальные уравнения и те же рецепты перехода к физическим объектам, как в классике, так же беспочвенно, как было беспочвенно требование свести оптику на механику. Более того, не только нет оснований считать такое сведе́ние возможным, но, наоборот, есть основание считать его, если бы оно даже было возможным, неудовлетворительным.

Итак, сам факт отличия квантовых законов от классических не требует оправдания. Отличие естественно и ни в коем случае не есть минус теории.

Совершенно другой вопрос: насколько отличаются и отличаются ли вообще общие принципы построения квантовой теории от тех принципов теоретико-познавательного характера, которые ле-

---

[1] [Процесс сведе́ния.]

[2] [Продолжающийся до бесконечности.]

жали в основе физических теорий классического периода? На этот вопрос, о котором я упомянул в самом начале, мне кажется, нужно ответить отрицательно.

Всякая физическая теория может быть расчленена на следующие этапы.

Прежде всего теория вводит, в зависимости от той или иной области физических явлений, к которым она относится, те или иные математические величины. Например, в механике такими величинами являются координаты, скорость, время и т. д.; в учении о теплоте — температура, количество тепла, энтропия и пр.; в электродинамике — электрическое и магнитное поля. Между этими величинами устанавливаются математические соотношения, например, в виде дифференциальных уравнений: механические уравнения Ньютона, электродинамические — Максвелла. Этот этап, который иногда сам по себе носит название физической теории, ею ни в коем случае не является, так как вообще не содержит никаких высказываний, относящихся к физическим явлениям.

Второй этап состоит в сопоставлении математических величин с физическими объектами. Это достигается тем, что для каждой величины дается определенный рецепт, указывающий, при помощи какого приема мы можем данному физическому объекту сопоставить то или иное численное значение этой величины. Мы называем это рецептом измерения данной физической величины. Наконец, те же рецепты служат для обратного перехода от новых значений величин, полученных из первых при помощи указанного выше формального математического аппарата, опять к физическим объектам.

Пример лучше всего пояснит сказанное. Мы измеряем, например, положение планеты в данный момент времени. Рецепт таков: мы отсчитываем соответственные деления лимба трубы, когда пересечение окулярных нитей совпадает с данной планетой, и одновременно положение стрелки часов, т. е. мы определяем значение угловых координат в данный момент времени. Затем мы подставляем полученные таким образом числа в соответственные уравнения в качестве начальных условий. Уравнения дают нам новые совместные значения координат и времени (t, φ, ψ). Высказывание теории заключается в утверждении, что, установив трубу соответственно с полученными из этих уравнений значениями делений кругов, мы увидим планету на пересечении нитей одновременно с прохождением стрелки часов через число на циферблате, равное t.

Только совокупность всех указанных этапов и составляет теорию.

Формализм, т. е. математические соотношения между символами, сами математические символы и рецепты сопоставления их с физическими объектами от теории к теории различны; но указан-

ные этапы и их совокупность характерны для всякой теории. Под эту структуру подходит и классика, и квантовая теория.

На одно, правда, принципиально мало существенное различие я хотел бы еще указать. Исторически в классике установление рецептов измерения значений величин обычно предшествовало установлению формализма, т. е. уравнений. Больше того, мы так сжились с этими рецептами издавна, что почти не замечаем их условности и принципиального значения. Поэтому-то их часто вовсе не причисляют к составным частям теории, что, конечно, как сказано, ошибочно. В противоположность этому в современной теоретической физике и особенно в квантовой теории рецептурная часть гораздо менее привычна и занимает большее место.

Более того, в современной теории квантов дело обстояло исторически так, что на основании некоторых довольно расплывчатых соображений установили сначала формализм — я имею в виду уравнение Шрёдингера — по отношению к некоторому символу, к так называемой функции ψ. И только после этого поставили себе вопрос: какой нужно дать рецепт для сопоставления ψ объектам природы? Как известно, современный рецепт этой связи был установлен не сразу, а только после некоторых проб, оказавшихся неудачными.

После этих предварительных замечаний, которые, как мне кажется, дают некоторую характеристику того положения, которое занимает квантовая теория как физическая теория вообще, без учета специфичных, присущих специально ей сторон, я хотел бы перейти к намеченным вначале проблемам, касающимся как раз тех сторон квантовой концепции, которые так резко отличают ее от всех предыдущих теоретических концепций.

Подойти к основным вопросам квантовой механики можно с различных сторон. Для нашей цели лучше всего обратиться — причина этого будет видна дальше — к историческому развитию квантовой теории, или, вернее, к той ее черте, которая впервые была выдвинута Планком и определила сразу же совершенно своеобразную ее структуру. Я имею в виду утверждение о дискретности возможных значений энергии данной механической системы (например, линейного осциллятора), могущей с точки зрения классической механики обладать непрерывной последовательностью энергетических состояний. Планк, как известно, впервые выдвинул квантовую гипотезу в вопросе о черном излучении; ею воспользовался Эйнштейн в проблеме удельных теплот твердого тела и, наконец, Бор — в вопросе о спектральных сериях. В этих трех фундаментальных для физики вопросах, абсолютно не укладывающихся в рамки классических воззрений, гипотеза о дискретности допустимых состояний механической системы впервые дала возможность внести ясность и получить удовлетворительное, во

всяком случае в принципиальном смысле, согласие между теорией и опытом.

С другой стороны, хорошо известно, что в основных проблемах, связанных с колебаниями, например с колебаниями струн, мембран и т. д., мы имеем дело с дискретной последовательностью возможных частот. Математический формализм этих явлений, вытекающий из ньютоновских уравнений движения, хорошо изучен, — это так называемые краевые задачи. Сводится дело на то, что физический объект — «частота» — сопоставляется с собственными значениями определенных дифференциальных операторов.

Дальнейший этап — математический факт, состоящий в том, что при известных «краевых» условиях собственные значения оператора составляют дискретную последовательность. Шрёдингер сопоставил удачу гипотезы о дискретности энергетических состояний систем микромира и наличие готового математического аппарата, естественным образом выделяющего дискретную последовательность чисел. Ему пришла в голову мысль воспользоваться с соответствующими изменениями этим аппаратом для создания общей теории микросистем. Его первые основные работы соответственно этому озаглавлены: «Die Quantisierung als Eigenwertproblem» [1].

В свете того, что было сказано выше о структуре физической теории, дело обстоит так. В математических соотношениях, лежащих в основе проблемы собственных значений, типично естественное выделение дискретной совокупности чисел. В классике с этим математическим аппаратом были связаны величины, относящиеся к проблемам колебаний макротел. Шрёдингер постулировал положение, что теми же (с некоторыми изменениями) соотношениями связаны и совершенно другие величины, а именно, величины, относящиеся к поведению микросистем.

В эти математические соотношения входит ряд величин: собственные значения операторов, собственные функции и т. д. Шрёдингер постулировал далее, что некоторый определенный линейный дифференциальный эрмитов оператор соответствует энергии, а его собственные значения соответствуют возможным ее численным значениям. Уже этого, т. е. *вида операторов*, для некоторых «моделей» атома было достаточно, чтобы выявилась плодотворность такого построения. При этом вопрос о том, какое же содержание имеют другие математические величины, оставался открытым. Дальнейшее развитие теории, коротко говоря, состояло в следующем: был введен постулат, согласно которому не только энергии, но и другим физическим величинам сопоставляются определенные эрмитовы операторы (координате q — умножение

---

[1] [Квантование как задача о собственных значениях.]

на q, импульсу — оператор $\frac{\hbar}{i}\frac{\partial}{\partial q}$ и т. д.) в сочетании с определением, что их собственные значения отвечают возможным значениям соответственной физической величины.

Конечно, спектр собственных значений как самой энергии, так и других величин, вообще говоря, не всегда дискретен. Есть случаи, когда и энергия может принимать непрерывную последовательность значений. Таким образом, произошло естественное расширение квантового формализма, которое и придало всей теории необходимую общность.

Следующий шаг состоял в постулате: состояние системы определяется некоторой нормированной к единице функцией от q и t, $\psi(q, t)$ [где q — спектр (непрерывный) оператора «координата», t— параметр], которая следующим образом связывается с физическими объектами. По определению, вероятность застать в момент времени t значение данной физической величины, равное i-му собственному значению соответствующего этой величине оператора, равняется квадрату коэффициента при i-ой собственной функции в разложении ψ по собственным функциям оператора.

Наконец, последний шаг, завершающий математическую часть теории, состоит в указании того дифференциального уравнения, которому подчиняется $\psi(q, t)$ как функция от q и t. Это уравнение — временно́е уравнение Шрёдингера — позволяет определить состояние системы в момент времени t, если задано состояние (т. е. ψ) в момент $t = 0$. Наряду с состояниями, определяемыми волновой функцией, так называемыми чистыми случаями, рассматриваются и другие совокупности — так называемые смеси. К этому формализму присоединяются рецепты для измерения значений соответственных величин в данный момент времени.

Таким образом, если сравнить структуру математической части квантовой теории с математической частью классической механики, то можно приблизительно сказать следующее. Сопоставление физических величин линейным операторам вместе с указанным выше толкованием их собственных значений и функций соответствует в классике тривиальному положению, сопоставляющему каждому значению физической величины определенное число.

Временно́е уравнение Шрёдингера соответствует в классике уравнениям движения, например уравнениям Ньютона.

Из изложенной здесь чисто схематически теории выводится уже чисто формальным путем ряд следствий, некоторые из которых имеют несомненно принципиальное значение. Одно из таких следствий — это соотношение неопределенностей Гейзенберга.

Из вышеизложенного ясно, что величины квантовой механики относятся не к индивидуальным случаям, а к совокупностям. Соотношение неопределенностей Гейзенберга показывает, что кван-

товомеханический формализм не противоречив только в том случае, если совокупности обладают некоторым свойством, а именно тем, что между стандартом импульса и стандартом координаты существует неравенство

$$\Delta p \cdot \Delta q \geqslant \frac{\hbar}{2}.$$

Иными словами, в теории отрицается содержательный смысл каких бы то ни было высказываний относительно таких совокупностей, в которых одновременно координата и импульс имеют определенное значение.

Классическая теоретическая механика в противоположность этому имеет дело преимущественно с индивидуальными явлениями, или, что в конце концов то же самое, с совокупностями, в которых обе величины — и q, и p — имеют одновременно вполне определенные значения.

Мы склонны придавать этому различию огромное значение (оно-то и заставляет нас говорить о новом физическом мировоззрении). Для пояснения я хотел бы сказать следующее. Конечно, и в макромеханике у нас нет возможности практически создать такую совокупность опытов, чтобы в каждом из них точно повторялись начальные значения координат и импульсов. Поэтому мы и не ожидаем, что при помощи уравнений движения мы сможем *точно* предсказать координаты и импульсы в какое-нибудь более позднее время. Но мы принимаем как данное, что принципиально нет препятствий к тому, чтобы сделать разбросы начальных значений координат и импульсов как угодно малыми. Все дело за техническими средствами. Формализм классической механики приурочен именно к этому предельному случаю и позволяет предсказать будущие значения координат и импульсов с той же точностью, какою обладают начальные значения, т. е. в пределе — с абсолютной точностью.

Утверждение возможности беспредельного уточнения начальных условий и q, и p и соответственного беспредельного уточнения предсказания будущих значений этих величин при помощи соответствующего математического аппарата — характерная и существенная черта классической механики.

Соотношение Гейзенберга, как мы только что видели, показывает, что квантовая теория отрицает принципиальную возможность такого уточнения начальных условий, так как согласно этому соотношению произведение стандартов q и p не может быть меньше конечной величины $\hbar/2$. Тем самым отрицается, конечно, и точное предсказание дальнейших значений.

Все это противоположение классической теории и квантовой заставляет нас говорить о новом физическом мировоззрении, и

нужно признать, что это отличие действительно весьма существенно. Вопрос другой, насколько прежние классические теории были действительно последовательны до конца. Но это отдельный вопрос, разбор которого нас завел бы слишком далеко.

Тесно связана с указанной неопределенностью еще одна характерная черта квантовой механики, которая обусловливает ее своеобразие и тоже чужда классике. Я имею в виду воздействие процесса измерения на измеряемый объект, которое в квантовой механике, в противоположность классике, не может быть сделано меньше вполне определенной величины и не поддается учету

Как бы там ни было, нужно признать, что квантовая механика настолько отличается от того, чем мы, во всяком случае, считали все существовавшие до сих пор физические теории, что мы говорим о новом мировоззрении. Нет ничего удивительного в том, что признание соотношения неопределенностей встречает сильное противодействие. Это выражается, между прочим, в стремлении считать эту сторону квантовой теории провизорной, а это стремление, в свою очередь, приводит естественным образом к следующему вопросу. Существуют ли основания утверждать, что нельзя дополнить или перестроить квантовую механику таким образом, чтобы сохранились те ее стороны, которые так блестяще оправдали себя в применении к истолкованию и предвидению физических явлений, и чтобы в то же время она не содержала соотношения неопределенностей?

На так поставленный вопрос трудно ответить, поскольку a priori нельзя сказать, что́ значит перестроить теорию. Я думаю, что на вопрос, поставленный в такой общей форме, ответа никто дать не может. Если иногда говорят, что Нейман доказал невозможность построения каузальной атомистической теории, то это неверно. Нейман сам никогда этого и не утверждал.

Но можно, сузив несколько вопрос, прийти к такой формулировке, которая еще достаточно обща, несомненно достаточно интересна и в то же время допускает точный ответ. Это и сделал, по существу, Нейман. Прийти к этой формулировке можно на основании следующих соображений. Несомненно и, как я думаю, ясно из предыдущего, что успех квантовой теории в первую очередь был обусловлен тем, что физическим величинам, в первую очередь энергии, были сопоставлены линейные операторы; ведь именно это и позволило применить математический аппарат задач о собственных значениях и т. д. Поэтому вполне целесообразно считать сопоставление физическим величинам линейных эрмитовых операторов тем основным моментом, которому обязана квантовая механика своей действенностью.

Это сопоставление ни в какой мере еще не предопределяло развития теории. Истолкование собственных значений операторов и

собственных функций, не говоря уже о форме уравнения движения, все эти дальнейшие постулаты — положения, которые составляют содержание специальной современной квантовой теории, — ни как не следуют из чрезвычайно общего первого постулата, допускающего, конечно, принципиально и другое развитие. Но, с другой стороны, если отказаться от первого постулата, то вообще ничего от современной квантовой механики не останется. Таким образом, мы его кладем в основу. Для исчерпывающей характеристики вопроса нужно предъявить еще некоторые дополнительные требования, и тогда мы получим ту постановку вопроса, которую дал Нейман и из которой я и хотел бы исходить.

Итак, схематически мы ставим, следуя Нейману, следующий вопрос:

Какие следствия могут быть выведены из предположения (с указанным ниже уточнением), что каждой физической величине соответствует линейный эрмитов оператор? Или еще так:

Можно ли построить теорию, в основе которой лежит (с некоторыми довольно существенными уточнениями) постулат, что каждой физической величине соответствует эрмитов оператор, и которая в то же время не содержала бы соотношения неопределенностей?

Насколько предопределены те черты квантовой механики, которые делают ее столь отличной от классики — соотношение неопределенностей и факт невозможности измерить что-нибудь без конечного воздействия,— тем постулатом, который собственно и определил ее успех и который в основном состоит в утверждении, что в формализме, наиболее подходящем для объяснения микроявлений, физическим величинам нужно сопоставлять линейные эрмитовы операторы и устанавливать между ними соотношения?

Нужно заметить для пояснения следующее: физическая величина *определена* и здесь тем, что предполагается данным рецепт для нахождения ее численных значений, рецепт для ее измерения. Но, в отличие от классики, формализм квантовой механики относится не к этим числам, а к представляющему величину оператору, с которым в силу определения эти числа приведены в строго сформулированное соответствие, например, являются его собственными значениями.

СЕМИНАР ПО НЕКОТОРЫМ ВОПРОСАМ ОПТИКИ И ТЕОРИИ КОЛЕБАНИЙ

(1939—1940 гг.)

О РАССЕЯНИИ СВЕТА

(20.XII 1939 г.)[1]

Еще совсем недавно весь комплекс вопросов, связанных с рассеянием света, был неясен. Обидно, что вам неизвестен весь генезис этого вопроса. Речь идет здесь об истории самого последнего времени, и актуальность этого генезиса еще очень велика. Все происходило на моих глазах. Несомненно, это одна из интереснейших страниц истории физики.

Что является стержнем, позволившим расшифровать весь комплекс явлений? Это — вопрос о тепловой энергии тела. Оказалось, что существует тесная связь между теорией теплоемкости и теорией рассеяния света. Рассеяние — оптическое проявление тех же процессов, которыми обусловлена теплоемкость тел. Понимание того, что теплоемкость и рассеяние — это одно и то же, было главным фактором в раскрытии всего вопроса. Раньше не приходило в голову, что эти вещи как-то связаны.

Как подошли к этому пониманию?

Вопрос о теплоемкости — гораздо более старый вопрос. Для газов все казалось давно выясненным. Относительно твердых тел 30 лет тому назад просто думали, что каждый атом колеблется около *своего* положения равновесия, с которым он упруго связан. Классическая физика рассуждала так: твердое тело имеет $3n$ степеней свободы; на каждую приходится кинетическая энергия $kT/2$ и потенциальная энергия $kT/2$, т. е. полная энергия kT.

---

[1] [Из материалов оптического семинара 1939/40 г. сохранились записки двух выступлений Л. И. Мандельштама, которые здесь и помещены. Одно из них, посвященное рассеянию света, записано и обработано Г. С. Гореликом. Второе — это две лекции об отклонениях от законов Френеля. Первая лекция записана и обработана С. М. Рытовым. Вторая лекция обработана Г. С. Гореликом по собственным запискам и затем несколько пополнена по запискам С. М. Рытова.]

Следовательно, энергия всего тела

$$U = 3nkT,$$

откуда

$$C_v = 3nk.$$

Итак, произведение атомного веса на удельную теплоемкость есть постоянная величина $C_v = 6$ *кал/град* (закон Дюлона и Пти). Действительно, опыт показывает, что для алюминия (атомный вес 27,1) $C_v = 5{,}5$, для золота (атомный вес 196) $C_v = 5{,}9$. Правда, такое хорошее совпадение получается не всегда. Уже давно было известно, что для тел с малым атомным весом дело обстоит не так хорошо.

Надо сказать, что Релей *никогда* не удовлетворялся этой теорией. Заведомо есть еще другие степени свободы, помимо тех, о которых шла речь. Об этом свидетельствуют оптические спектры. Между тем «степени свободы не взвешиваются, а *считаются*», и *каждая* степень свободы должна иметь кинетическую энергию $kT/2$. Это как раз и плохо.

Но что же следует из этих вещей по отношению к рассеянию света? До начала XX в. вообще мало занимались рассеянием. Думали, что все дело в «загрязнениях». Еще не так давно считали, что путем тщательной очистки жидкость можно сделать «оптически пустой».

Эйнштейн сделал первый шаг в вопросе о теплоемкости. Он увидел, что воззрения Планка должны и здесь, [т. е. не только в теории излучения] что-то изменить [1]. Закон равномерного распределения энергии по степеням свободы *несправедлив*, на каждую колебательную степень свободы приходится средняя энергия не kT, а

$$u = \frac{h\nu}{e^{h\nu/kT} - 1}.$$

При $T \to 0$ это дает $u \to 0$; при $\nu \to \infty$ также $u \to 0$. Теплоемкость должна вычисляться по формуле

$$C_v = \frac{d}{dT}(3nu) = f(\nu),$$

т. е. она оказывается теперь функцией частоты собственных колебаний атомов. Вот это главное, и это было с облегчением встречено физиками. (В сущности, для вычисления полной внутренней

---

[1] [*A. Einstein.* Ann. Physik, 22, 180, 800, 1907.]

энергии тела надо не множить u на $3n$, а брать сумму

$$\sum_i \frac{h\nu_i}{e^{h\nu_i/kT}-1},$$

распространенную на все частоты собственных колебаний. Из оптики известно, однако, что для дальнейших членов ν_i очень велики и, таким образом, эти члены отпадают).

Релей не принимал этого решения вопроса, он не мог с этим сжиться.

О рассеянии света пока не говорится ни слова, и на данном этапе это и не удивительно. Интересно, как дело пошло дальше.

Тот же Эйнштейн взялся за рассеяние [1]. Он был тогда в расцвете сил, во всем, за что он брался, он делал фундаментальные вещи.

Эйнштейн говорит следующее. Из-за флуктуаций плотность тела является функцией точки, $\rho = \rho(x, y, z)$. Можно разложить $\rho(x, y, z)$ в ряд Фурье, изучить статистически амплитуды разложения и найти по ним рассеяние. Удивительно, что при этом не говорится *ни слова* о теплоемкости. У Эйнштейна ρ задано статически, никакой динамики нет.

Два года спустя к теплоемкостям возвращается Дебай [2], у него новая, очень красивая идея. Верно то, что твердое тело имеет $3n$ степеней свободы. Но это степени свободы *связанных* осцилляторов. Некоторые степени свободы мы хорошо знаем: это — упругие волны. Они тоже являются собственными колебаниями системы связанных осцилляторов. Предположим, что всем $3n$ степеням свободы соответствуют упругие колебания; тогда существуют $3n$ отдельных упругих волн.

Здесь в первый раз высказывается мысль, что тепловая энергия есть энергия упругих волн. Но о рассеянии — опять *ни слова*.

Однако к нему подошли очень близко. Оставался только последний шаг. Надо было сказать: «формальные волны» Эйнштейна, на которых происходит рассеяние, *это и есть* дебаевские волны, определяющие энергию теплового движения тела.

Теперь это кажется очень простым, но в то время такие люди, как Эйнштейн и Дебай, не видели связи. В последующие годы удалось показать, что это одно и то же, но досталось это не так легко [3].

---

[1] [*A. Einstein*. Ann. Physik, 33, 1275, 1910.]
[2] [*P. Debye*. Ann. Phys., 39, 789, 1912.]
[3] [См. том I, статьи 22, 27 и 29.]

ОТКЛОНЕНИЯ ОТ ЗАКОНОВ ФРЕНЕЛЯ

ПЕРВАЯ ЛЕКЦИЯ

(26.II 1940 г.)

Сегодня и в следующий раз мы рассмотрим вопрос об отражении волн на границе двух сред. Здесь есть целый ряд тонкостей и принципиальных моментов, на которые зачастую не обращают достаточного внимания.

Большой интерес имеют вопросы, связанные с существованием и поведением так называемых *поверхностных* волн. С ними приходится иметь дело, в частности, при рассмотрении упругих волн на границе двух тел. Сначала ими занимался Релей, потом ряд других ученых. Изучение этих волн важно и с практической точки зрения — для сейсмологии. В электродинамике, когда обратились к распространению радиоволн над поверхностью Земли, всплыл аналогичный вопрос. На первых порах здесь руководились соображениями, почерпнутыми именно из теории упругости, и отсюда Ценнек пришел к своей теории поверхностных волн. Однако это было не к пользе для дела. В чем различие этих двух случаев — упругих и электромагнитных поверхностных волн, что характерно для каждого, это мы тоже рассмотрим на дальнейших занятиях семинара. Затем я думал возобновить курс квантовой механики.

Сегодня мы займемся вопросом, который на первый взгляд кажется совершенно элементарным.

Имеются две оптически прозрачные однородные среды с плоской (для начала) границей раздела. Речь идет об отражении и преломлении плоской волны на этой границе — вопрос, рассматриваемый во всех учебниках. Но в нем есть два момента, которые нужно, по-моему, углубить.

Во-первых, всегда ли нам достаточно знать решение задачи только для плоской границы? Одно из самых идеальных осуществлений плоской границы — это свободная поверхность жидкости. Но согласно взглядам современной статистической физики поверхность жидкости *принципиально* является волнистой, очень мало, но неизбежно. Как отражается свет от такой «плоской» поверхности? Этот вопрос удовлетворительно (хотя и не исчерпывающим образом) разобрал Релей. Первое, что нужно здесь выяснить,— это особенности отражения от волнистой поверхности, но сегодня мы этого делать не будем.

Во-вторых, даже в обычном случае идеально плоской границы раздела вопрос совсем не так прост. В учебниках его изложение

схематизировано. Между тем опыт обнаруживает расхождения с теорией. Вопросы здесь есть, но мы можем только указать на них, так как ответов пока не дано.

Мы требуем от электромагнитной теории света, чтобы она дала нам законы отражения и преломления и выражения для интенсивностей отраженной и преломленной волн. Казалось бы, формулы Френеля полностью решают дело. В частности, для коэффициентов отражения (которое только и будет нас сейчас интересовать) получаются выражения

$$R_p = \frac{\operatorname{tg}(\theta - \varphi)}{\operatorname{tg}(\theta + \varphi)}, \qquad R_s = -\frac{\sin(\theta - \varphi)}{\sin(\theta + \varphi)},$$

где R_p относится к волне, у которой колебания происходят в плоскости падения, а R_s — к волне, у которой они перпендикулярны к этой плоскости, и где угол падения θ и угол преломления φ связаны законом преломления

$$\sin\theta = n \sin\varphi.$$

Первый, кто дал эти формулы, был Френель. Это было в 1823 г., когда все безраздельно стояли еще на точке зрения упругого эфира. Сказать, что Френель *вывел* эти формулы, было бы, пожалуй, слишком сильно. Скорее он их угадал, как он всегда угадывал: путем не особенно верных рассуждений он приходил к правильным результатам. Это был гений особого рода.

Очень много занимался вопросом об отражении и преломлении Коши. Он выводил законы этих явлений из граничных условий на поверхности раздела, но действовал при этом как математик, формально. Между тем, если рассматривать эфир как упругую среду, то физически необходимы шесть граничных условий, выражающих отсутствие разрывов на поверхности раздела для трех смещений и трех напряжений. Этим шести условиям можно удовлетворить, только допуская наличие в среде и продольных, и поперечных волн, т. е. считая упругую среду твердым телом. Таким образом, световой эфир должен быть твердым телом, в нем должны распространяться и поперечные, и продольные волны. Опыт же говорил, что световые колебания всегда поперечны. Было много попыток выйти из этого противоречия. Были, например, попытки считать эфир абсолютно несжимаемым. Продольные волны тогда невозможны, но зато возникают противоречия в объяснении других оптических явлений (распространение света в кристаллах).

В электромагнитной теории Максвелла все эти трудности отпали. Она требует выполнения только *четырех* граничных условий — непрерывности тангенциальных компонент напряженностей и нормальных компонент индукций. Формулы Френеля получают-

ся из нее естественно, и никаких продольных волн в теории не фигурирует. Исторически интересно, что в своем первом трактате Максвелл не касается вопроса об отражении. Уиттекер замечает по этому поводу, что Максвелл был еще слишком сильно связан тогда представлениями теории упругости и думал, что условий должно быть не четыре, а *шесть*. Гельмгольц первый показал, что в электромагнитной теории достаточно четырех условий, потом это подробно разобрал в своей докторской диссертации Г. А. Лоренц.

Но ограничиваясь четырьмя условиями и получая из них формулы Френеля, мы еще не избавляем себя от неожиданных парадоксов. Известно, что можно так подобрать толщину плоскопараллельной пластинки из прозрачного материала, чтобы она ничего не отражала. Сколь угодно толстая стопа из таких пластинок тоже не будет ничего отражать. В пределе, увеличивая число пластинок, мы можем получить полупространство, заполненное нашим прозрачным материалом и не дающее отражения, в то время как формулы Френеля, выводимые как раз для границы такого полупространства, дают вполне определенное отражение с коэффициентами R_p и R_s. В чем же дело?

Дело в том, что обычный вывод формул Френеля не полон. Мы с самого начала не вводим в рассмотрение волн, идущих во второй среде *снизу*, а берем в этой среде только одну (преломленную) волну. Это ниоткуда не следует. Во всяком случае, это не содержится в четырех граничных условиях, которые вполне допускают наличие *четырех* волн — двух в первой среде и двух во второй

Здесь есть и еще один, гораздо более интересный момент. Во всех теоретических выводах закона преломления последний берется только в одном из возможных его видов, а именно, преломленная волна уходит от границы раздела во вторую среду и нормаль этой волны лежит по другую сторону от нормали к границе, чем у падающей волны (рис. 1, *а*). Но условия на границе, не зависящие от времени, допускают и другой случай, изображенный на рис. 1, *б*. Почему этот случай отбрасывается? Откуда мы знаем, что преломленная волна идет вниз, а не вверх? Это тоже *привносится* в решение задачи извне и связано с интуитивным убеждением, что групповая скорость направлена так же, как и фазовая [1].

Но все это до некоторой степени формальные моменты. Кроме них есть еще и физические интересные вопросы, к которым мы теперь и перейдем.

---

[1] [Подробно этот вопрос был рассмотрен Л. И. Мандельштамом в лекциях «Некоторые вопросы теории колебаний» (стр. 434—435) и в статье «Групповая скорость в кристаллической решетке» (том II, статья 52).]

Мы имеем формулы Френеля и можем испытать их экспериментально. Для угла падения $\theta = 0$ коэффициент отражения равен

$$R = \frac{n-1}{n+1}.$$

Сделаем опыт и профотометрируем отраженный свет. Такого рода исследованиями занимался Релей. Он показал, что коэффициент отражения очень нечувствителен к отклонениям от нормального падения. Так, например, для воды при $\theta = 10°$ угол преломления $\varphi = 7°29'$, а коэффициент отражения только на 0,4% превышает значение при нормальном падении. Значит, малые отклонения пучка от нормали и от строгой параллельности не могут служить объяснением больших отклонений экспериментальных значений R от теоретических. Между тем, при точности измерений порядка 0,5—1% Релей получил уклонения от формул Френеля немного более 1%. Он не сделал отсюда выводов, но все время относился к формулам Френеля с некоторым недоверием.

Для суждения о согласии или несогласии формул Френеля с опытом особенно пригодны так называемые нулевые эффекты, т. е. исчезновение какого-либо явления, например отражения, при определенных условиях. На основании формул Френеля можно ожидать двух таких эффектов при отражении.

Во-первых, на границе двух тел, показатели преломления которых для какой-либо длины волны λ одинаковы (скажем, кронглас

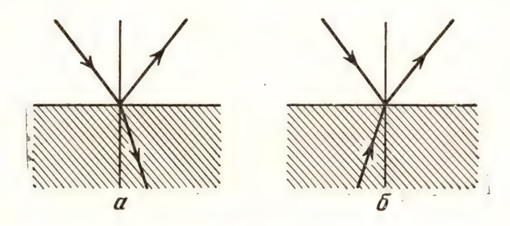


Рис. 1

и смесь бензола с сероуглеродом), отражение света этой длины волны должно исчезать. Такого рода опытами занимался Жамен и установил, что отражение все-таки остается, и притом довольно сильное. Конечно, если поверхность стекла загрязнена, то тогда ясно, что нулевого эффекта быть не должно. Но оказалось, что свести отражение на нет никакими химическими способами очистки не удается. При обработке поверхности плавиковой кислотой отражение резко и заметно падает, однако некоторая величина остается и она *постоянна*. В чем же здесь дело? Несомненно, это

первое серьезное нарушение требований теории. Можно, конечно, сослаться на то, что иметь дело с твердыми телами в таких опытах, наверно, неприятно и что лучше взять две жидкости (подобные смеси — с одинаковым показателем преломления, но разной плотностью — приготовить можно). Хотя сегодня и нет таких опытов, но тем не менее можно сказать с полной уверенностью: остаточное отражение существует.

Во-вторых, формулы Френеля предсказывают нулевой эффект (отсутствие отражения) при падении света с p-поляризацией под углом Брюстера, т. е. при

$$\operatorname{tg}\theta = n, \text{ или } \theta + \varphi = \frac{\pi}{2}.$$

В этом случае $R_p = 0$. Отклонения от этого закона также были замечены еще Жаменом. Он нашел, что никогда не удается получить нуль, а достигается только минимум R_p. По формулам Френеля, при переходе через угол Брюстера R_p меняет знак, т. е. фаза отраженной волны должна скачком меняться на 180°. Непрерывность соблюдается благодаря тому, что в этот момент интенсивность проходит через нуль. Если же интенсивность не достигает нуля, то изменение фазы должно происходить непрерывно. Жамен это и обнаружил: при угле Брюстера сдвиг фазы равнялся 90°. Если записать R_p в виде

$$R_p = \rho e^{i\delta},$$

то по величине ρ при $\operatorname{tg}\theta = n$ можно судить об уклонении от формул Френеля. Для воды Жамен получил $\rho = 0{,}0058$ ($\delta = 90°$). Для твердых тел (стекло) отклонения еще больше, но существенно, что они вообще есть и наблюдаются у всех тел.

Первый, кто пытался дать теорию этих аномалий, был Друде. Его теория, будь она верна, разрешила бы все вопросы. Он предположил, что на поверхности имеется тонкая пленка, не грязь, а переходный слой, толщина которого много меньше длины волны. В этом случае, если даже показатель преломления в таком слое непрерывно меняется по его толщине, очень легко рассчитывать, как ρ зависит от угла падения, толщины пленки и ее показателя преломления[1]. Получается, что отражение есть и при угле Брюстера. Друде приводит такой пример: у тяжелого флинта $n = 1{,}75$. Если предположить, что у пленки n равно среднему арифметическому показателей стекла и воздуха, то достаточно толщины пленки $d = 0{,}0017\,\lambda$, чтобы получить при угле Брюстера $\rho = 0{,}003$. Таким образом, речь идет о слоях толщиной порядка десятка межатомных расстояний.

---

[1] [См., например, *Борн*. Оптика. Харьков — Киев, ОНТИ, 1937, стр. 52; *П. Друде*. Оптика. М.— Л., ОНТИ, 1935, стр. 238.]

Казалось бы, все в порядке и объяснение найдено. В пользу этой теории говорило и то обстоятельство, что у кристаллов на свежем разломе наблюдалось заметное и быстрое увеличение ρ. Но Релей все же таким объяснением не удовлетворился. Он поставил очень простые и в то же время чрезвычайно тонкие опыты, пользуясь поверхностью воды. По сути дела, вся химия пленок на поверхности жидкости идет от Релея.

Пусть дело заключается в пленках, говорит Релей. Будем максимально очищать поверхность воды и затем специально покроем ее пленкой. При чистых веществах эта пленка обязательно получается мономолекулярной: ее просто не удается сделать толщиной в 2—3 молекулярных слоя, так как при избытке наносимого вещества получается слой толщиной в одну молекулу, а остальное собирается в капли. Подобное загрязнение поверхности само по себе ничтожно, но на оптических ее свойствах оно уже существенно сказывается. Релей получил значения ρ для воды того же порядка, что и Жамен. В процессе очистки абсолютная величина ρ уменьшалась (ρ может быть как больше, так и меньше нуля, в зависимости от показателя преломления пленки). С усилением очистки поверхности Релей дошел до $\rho = 0$. Таким образом, все как будто бы получилось хорошо. Но тут он обнаружил, что при дальнейшей очистке ρ снова появляется, переменив знак. Пользуясь в качестве источника света солнцем, Релей измерил значение $\rho = 0{,}0004$. От чего может получиться такой эффект? Очевидно, обращение ρ в нуль не есть результат окончательной очистки, а представляет собой эффект *компенсации*. Другими словами, имеется какая-то причина существования у воды и вообще у всех веществ остаточного ρ, которая не сводится к наличию загрязнений.

Какие же факторы могут здесь оказать влияние?

Удовлетворительного объяснения нет, но, мне кажется, все же можно указать на то, что является ответственным за описанные явления. Дело, по-видимому, в том, что при описании «нулевых» эффектов среду уже нельзя считать однородным континуумом.

ВТОРАЯ ЛЕКЦИЯ

(20.III 1940 г.)

Законы Френеля грубо передают то, что наблюдается на опыте в случае отражения и преломления плоской световой волны на плоской поверхности раздела. Но, как мы подробно разобрали в прошлый раз, в действительности имеются существенные — не по величине, а по природе — отклонения от этих законов. Особенно отчетливо они проявляются в двух случаях. Согласно законам Френеля:

1) если $n_1 = n_2$, то не должно быть отражения;

2) если $\operatorname{tg}\theta = n$, то угол падения θ есть угол полной поляризации; компонента электрического вектора, лежащая в плоскости падения, в отраженном свете отсутствует.

Оба эти явления в действительности *не* наблюдаются. При $n_1 = n_2$ отражение все-таки есть. При $\operatorname{tg}\theta = n$ все-таки есть слабое отражение компоненты E, лежащей в плоскости падения.

Это — слабые эффекты, но они вызваны *не* загрязнениями. В их основе лежит что-то более принципиальное. Исследование этих эффектов наталкивается на большие трудности эксперимента: малейший слой постороннего вещества уже объясняет аномалии нужного порядка величины.

В чем здесь дело, мы не знаем. До сих пор вопрос остается открытым. Я укажу лишь на некоторые возможности объяснения. Они не просчитаны, так что уверенно ничего сказать нельзя, но, может быть, можно так подойти. Я думаю, что дело в молекулярной структуре, что континуум не может всего охватить и что именно *этот* момент надо учесть.

Начну с закона Брюстера.

Отраженный свет — это вторичное излучение. Излучение, обусловливающее отражение, «должно было бы» пойти по направлению отраженного луча. Но при угле Брюстера $\left(\theta + \varphi = \frac{\pi}{2}\right)$ оси осцилляторов направлены как раз по этому лучу, а вдоль оси осциллятор ничего не излучает.

Это объяснение было хорошо 20—25 лет тому назад. Сегодня это уже не так, ибо сразу же ставится вопрос, который делает это объяснение непригодным.

Мы считаем, что оси осцилляторов направлены по электрическому полю преломленной волны. Это хорошо, если осциллятор изотропен (я не говорю уже о том, что мы все делаем по классике. Кванты дадут здесь то же самое). Но сегодня мы знаем (хотя бы из деполяризации рассеянного света), что молекулы — не изотропные

осцилляторы. Поэтому под углом Брюстера нельзя ожидать полной темноты.

Но это все же не меняет дела. Каков бы ни был осциллятор, компоненты поляризации, перпендикулярные электрическому полю преломленной волны, излучаются *некогерентно*. Таким образом, наряду с когерентным излучением есть некогерентное, а значит, *и по этой модели* регулярного отражения под углом Брюстера не будет. Будет рассеяние света, но о рассеянии формулы Френеля вообще ничего не знают.

Наблюдаемое явление — регулярное отражение под углом Брюстера — может быть объяснено, если предположить, что вблизи поверхности молекулы ориентированы не беспорядочно, а имеют некоторое преимущественное направление, и тогда их анизотропия уже не дает *в среднем* того же, что получается для изотропных осцилляторов. Это не противоречит тому, что мы знаем о поверхности жидкости. Если это так (а так оно или нет, я не знаю), то должна быть связь между деполяризацией при рассеянии и отклонениями от закона Брюстера. Повторяю, вопрос остается открытым и теоретически, и экспериментально.

Есть и другое возможное объяснение. Показатель преломления внутри вещества обусловлен интенсивностью вынужденных колебаний осцилляторов. Но их амплитуда зависит от взаимодействия диполей, которое дает известную лорентц-лоренцовскую поправку в коэффициенте преломления. Обычный способ вычисления этой поправки: выделяется сфера и определяют отдельно поле диполей, находящихся вне сферы, и диполей, находящихся внутри нее. Хотя этот способ и мало убедителен, но он приводит к правильному результату. Однако на поверхности молекулы находятся в особых условиях, и этот способ с вырезанием сферы наверняка непригоден. Может быть, дело в этом, но вопрос открыт.

Вот еще третье возможное объяснение. Из-за дискретности вещества нужно на самом деле применять при вычислении электрического поля, создаваемого осцилляторами, не интегрирование, а суммирование. Расстояние a между осцилляторами очень мало по сравнению с длиной волны λ, так что $\frac{a}{\lambda} \ll 1$. В нулевом приближении, т. е. с точностью до $\left(\frac{a}{\lambda}\right)^0$, получатся формулы Френеля. Не может ли быть, что следующее приближение, т. е. члены порядка $\frac{a}{\lambda}$, объясняют отступления от законов Френеля? Все три перечисленных эффекта малы, но каждый может служить причиной наблюдаемых отклонений. Может быть, играют роль все три эффекта, может быть, два из них или только один. Вопрос этот хотя и трудный, но интересный, и пока для его решения еще ничего не сделано.

Были и трафаретные попытки объяснения, которые не дали ничего интересного. Укажу на две из них.

1. Предположим, что есть отличие не только между диэлектрическими постоянными обеих сред, но и между их магнитными проницаемостями, хотя и очень малое. Тогда для отсутствия отражения действительно недостаточно, чтобы было $n_1 = n_2$. Отражение будет, несмотря на равенство скоростей распространения в обеих средах. Чтобы отсутствовало отражение, должно быть не только $\varepsilon_1 = \varepsilon_2$, но и $\mu_1 = \mu_2$. Что касается угла Брюстера, то при несовпадении μ он остается, но несколько смещен. Опыт же говорит, что угла полной поляризации *нет*.

2. Среды имеют различную дисперсию, так что $n_1 = n_2$ выполняется только для какой-то одной волны. Между тем в са́мом монохроматическом свете всегда имеется некоторый конечный интервал для волн. Оказывается, однако, что обычная теория дисперсии не приводит здесь к тому, что нужно.

Вы видите, что ничего положительного я сказать не могу. Но одно несомненно: школьная трактовка слишком схематична. В природе дело обычно обстоит сложнее, и мы имеем здесь простой пример этого [1].

---

[1] [Вторая часть лекции была посвящена вопросу об отражении от шероховатой поверхности. Л. И. Мандельштам коротко рассказал теорию вопроса (см. том I, статью 19), после чего Г. С. Ландсберг сделал доклад об экспериментальных исследованиях рассеяния света поверхностью жидкости. См. *Ф. С. Барышанская.* ЖЭТФ, 7, вып. 1, 51, 1937.]

ЛЕКЦИИ
ПО НЕКОТОРЫМ ВОПРОСАМ ТЕОРИИ КОЛЕБАНИЙ
(1944 г.)

ПЕРВАЯ ЛЕКЦИЯ

(14.III 1944 г.)[1]

Приступая к нашим беседам, я хотел бы сегодня остановиться на некоторых общих вопросах, высказать некоторые соображения о цели этих бесед и о том материале, которым я буду пользоваться. Я хочу подчеркнуть заранее, что мы не будем рассматривать какие-либо сложные проблемы или специальные задачи. Наши беседы будут элементарны и будут касаться хорошо известных вещей, но мы постараемся посмотреть на эти известные вещи с несколько более общей точки зрения, чем это обычно делается.

В настоящее время вряд ли есть необходимость говорить о той громадной роли, которую играют колебания как в физике, так и в технике. Возьмите акустику, радио, оптику. Основа их именно колебательная. Не удивительно, что теперь, когда колебания приобрели такое большое значение и в технике, учение о колебаниях выделили как особую область. Каковы же те признаки, по которым выделяется учение о колебаниях? Присмотревшись, мы видим, что они принципиально отличны от тех, по которым делят физику на оптику, акустику и т. д. Это последнее деление производится, очевидно, по признаку физических явлений, которые мы одинаково воспринимаем. С электричеством и магнетизмом дело обстоит несколько сложнее (у нас нет непосредственного восприятия этих явлений), но я не буду на этом задерживаться. С колебаниями дело обстоит принципиально иначе: мы выделяем их не по физическому содержанию нашего восприятия, а по общности метода или

[1] [Материалами при обработке этих лекций послужили: 1) записки С. М. Рытова лекций 1 и 2, 2) записки С. П. Стрелкова лекций 1—4, 3) весьма лаконичные заметки И. М. Франка. Кроме того, были использованы чрезвычайно сжатые конспекты, составленные самим Л. И. Мандельштамом при подготовке к лекциям. Обработка указанного материала произведена С. М. Рытовым и С. П. Стрелковым.]

подхода к изучению, по общности формы закономерностей, независимо от физического содержания, или, вернее, при крайне разнородном физическом содержании. Здесь охвачены и акустические, и механические, и электрические, и оптические явления, которые чрезвычайно различны для нашего восприятия.

Именно это обстоятельство придает учению о колебаниях его громадное значение, его интерес. Изучая одну область, вы получаете тем самым интуицию и знания в совсем другой области. Вы получаете возможность проводить далеко идущие аналогии; темные места, скажем в оптике, освещаются, как прожектором, при изучении колебаний в механике и т. д. Таким образом, теория колебаний объединяет, обобщает различные области физики. Это положение можно пояснить следующим примером. Каждая из областей физики — оптика, механика, акустика — говорит на своем «национальном» языке. Но есть «интернациональный» язык, и это — язык теории колебаний. Она вырабатывает свои специфические понятия, свои методы, свой универсальный язык.

В наших беседах — это их цель — я хотел бы сделать ударение именно на этой обобщающей стороне. Я хотел бы, во-первых, обсудить некоторые из таких общих положений и, во-вторых, показать их действенность в возможно более разнообразных областях (нашей) науки. Обычно, излагая тот или иной предмет, мы стараемся дать конкретный материал, дать соответствующий математический аппарат, научить пользоваться этим аппаратом. С другой стороны, в оптике нас интересуют специфические оптические вопросы, в акустике — акустические и т. д. В результате получается разрозненность, за деревьями не видно леса. Это, конечно, естественно. Художник-специалист изучает на картине, как надо класть краски, как работать кистью и т. п. Но, для того чтобы получить общее впечатление, надо отойти от картины. Детали при этом теряются, но зато приобретается нечто другое. Мы видим тогда, как входят понятия в мировоззрение физика.

То, что я хотел изложить сегодня, можно назвать, пожалуй, комментированным оглавлением наших бесед. Я совсем кратко перечислю в качестве примера несколько вопросов, наглядно показывающих универсальность и мощь теории колебаний. Сейчас я их буду только называть. Разбор их, т. е. части из них, и составит предмет наших бесед. Как вы увидите, в этих примерах имеется пестрота, но она не случайна, а предумышленна.

Первое, на чем мы остановимся,— это некоторые вопросы кинематики колебаний, некоторые универсальные вещи, касающиеся формы колебаний.

На спокойной поверхности воды можно создать синусообразную волну

$$y = \alpha \cos (\omega t - kx),$$

где $\omega = \frac{2\pi}{T}$ — частота, соответствующая периоду T, а $k = \frac{2\pi}{\lambda}$ — волновое число, соответствующее длине волны λ. Вообще говоря, под скоростью волны понимают величину $\frac{\omega}{k}$ — скорость перемещения гребня.

В середине прошлого века Стокс заметил, что конечная группа синусообразных волн перемещается со скоростью, существенно отличной от фазовой скорости $\frac{\omega}{k}$ (Релей назвал эту скорость групповой). Сегодня понятие групповой скорости играет существенную роль. Она может совпадать с фазовой скоростью, но важно то, что бывают случаи, когда групповая скорость отлична от фазовой, а именно — всякий раз, когда фазовая скорость зависит от ω (или λ).

Для звуковых волн в воздухе практически нет зависимости $\frac{\omega}{k}$ от ω, и поэтому их групповая скорость равна фазовой. Поэтому музыка не искажается, если мы слушаем оркестр издалека. Будь это не так, то каждый аккорд размазался бы и получилась бы невероятная какофония. Для света положение иное. В вакууме тоже нет дисперсии, но в телах — воде, стекле и т. д. — это уже не так.

Замечание Стокса прошло незамеченным, хотя понимающие люди намотали его себе на ус.

В это время был другой вопрос — в оптике. До 50-х годов шла борьба сторонников эмиссионной теории Ньютона против волновой теории Юнга — Френеля. Все согласились, однако, в том, что есть один решающий опыт. При вхождении в более плотную среду световой луч приближается к нормали. С точки зрения корпускулярной теории это происходит потому, что скорость световых корпускул в среде больше, чем в воздухе. Волновая же теория требовала как раз обратного соотношения. Таким образом, оставалось только прямым опытом определить скорость в среде и это должно было решить вопрос. В 1850 г. Фуко и Физо произвели такой опыт, и оказалось, что в воде скорость меньше, чем в воздухе. Тем самым, как полагали, вопрос однозначно был решен в пользу волновой теории.

Но в конце прошлого века Майкельсон определил показатель преломления в сероуглероде для зеленой линии ртути. Он получил (из преломления) значение 1,64, в то время как прямые измерения скорости света в сероуглероде дали 1,77. Релей сразу же вспомнил замечание Стокса и указал, в чем здесь дело. Легко показать, что из преломления определяется отношение c к фазовой скорости, а все прямые методы определения скорости дают групповую скорость, отличную от фазовой при наличии дисперсии. Нет

скорости света. Есть много скоростей света, и одна из них — групповая.

Этот эпизод чрезвычайно существен. Существенно то, что нет понятия скорости при волновом движении, скорости *вообще*. Надо указать, *какой опыт* вы ставите, и только тогда определяется, о какой скорости идет речь. Это было ново и неожиданно. А между тем без этого нельзя было бы строить ни теорию относительности, ни волновую механику. Ведь на непонимании вопроса о скорости света как раз и строились некоторые возражения против специальной теории относительности в 1905 г. Говорили, что при аномальной дисперсии бывает фазовая скорость больше c. Но одно из основных положений теории относительности состоит в том, что *никакой предмет* или *сигнал* не может иметь скорость больше c. Если найти такой сигнал, то теория относительности пала бы. Теперь ответ на это прост: синусообразная волна, бесконечная в обе стороны, не может передать сигнал. Сигнал — это движение *признака* (увеличение и уменьшение амплитуды и т. п.). Но тогда это групповая скорость, а она меньше c.

Сложнее и интереснее обстоит дело в волновой механике. Там, если употребить грубую формулировку (я всегда против нее, но сейчас она для нас достаточна), свет должен быть и частицей, и волной, а скорость должна быть и больше, и меньше c. Это казалось сумасшествием, но теперь волновая механика решает этот вопрос совсем просто: электрон или вообще какая-либо частица — это группа волн и движется с групповой скоростью, а сами волны — с фазовой. Одна может возрастать, в то время как другая убывает, и т. д.

Удивительно, как трудно все это входит в плоть и кровь. Ведь теперь нам ясно, что Физо и Фуко ничего не доказали своими измерениями. Они мерили групповую скорость света, а ведь для доказательства правильности волновой теории надо было измерять *фазовую* скорость. Тогда этого не знали и, может быть, было счастьем для физики, что этот опыт почитался решающим, а его результат закрыл рот сторонникам эмиссионной теории. В физике, в науке вообще, неоднократно бывали такие случаи. Если бы знали, что при нагревании тело делается более весомым, то не высказали бы закона сохранения вещества, сыгравшего такую колоссальную роль в физике и технике.

Плохо, однако, когда современные физики не понимают сути дела. В «Истории света» Ронки (Ronchi) измерения Физо и Фуко по-прежнему объявляются решающим доказательством волновой теории: «В 1850 г. Физо и Фуко прямыми измерениями показали, что скорость света в вакууме больше, чем в материи. Отныне никто не мог думать о корпускулярной теории». Он не понимает, что (во всяком случае, логически) эти опыты ничего не доказывают.

Позвольте мне совсем коротко, в виде сухого перечисления, закончить вопросы кинематики.

Вы знаете, какую огромную роль сыграл принцип Допплера. В исследовании строения Вселенной он по сей день является одним из ведущих принципов. Но, вероятно, немногие знают, что когда капитан корабля для ослабления качки меняет курс или скорость, то он тоже исходит из принципа Допплера. Есть и другие вопросы, также «неожиданно» связанные с этим принципом.

Чисто кинематическими являются и вопросы интерференции. Сюда относятся не только замечательные цвета тонких пластинок и мыльных пузырей. Интерференцией объясняются фединги, или замирания, в радиоприеме. Без глубокого знания вопросов интерференции нельзя было бы объяснить синеву неба. Этот вопрос представляет собой другую сторону интерференции и, как это ни странно, связан с задачей о пьяном человеке. Человек этот настолько пьян, что каждый его шаг никак не связан с предыдущим шагом. Задача состоит в определении вероятности того, что за n шагов он уйдет настолько-то от исходной точки. Это задача о броуновском движении, но она же связана и с вопросом о синеве неба.

С точки зрения теории колебаний беспроволочная телефония и комбинационное рассеяние света — одно и то же. Это модуляция. Звук — в радио, колебания атомов — в комбинационном рассеянии. В 1928 г. я был в командировке за границей и встретился с одним из лучших мировых инженеров по радио. Он меня спросил: «Что это все говорят — комбинационное рассеяние. Не можете ли вы мне объяснить, что это такое?» Я сказал ему, что это модуляция. И для него, хорошо владеющего идеями колебаний, этого намека оказалось достаточно, чтобы он потом совершенно ясно представлял себе, что такое комбинационное рассеяние.

Это все вопросы кинематики колебаний. Я перейду теперь к другим очень важным вопросам. Речь идет о существовании *колебательных систем*, т. е. систем, устроенных так, что им присущи некоторые собственные колебания. Все вы знаете такие системы, как маятник и колебательный контур, и знаете, что это с колебательной точки зрения одно и то же. Теперь все это тривиально, но замечательно именно то, что оно тривиально. Существует, однако, множество других замечательных систем.

Земля покрыта водой. Каков период колебаний воды океанов, покрывающих Землю? Уже Ньютон знал, что это вопрос кардинальный для геофизики. Если представить себе, что вся вода находится в довольно широком канале, прорытом по экватору и имеющем глубину около 3 *км*, то период оказывается равным 30 часам. Таким образом, вода на Земле — это большой маятник с периодом в 30 часов. Не всегда представляешь себе, что это такое. На выставке

в Америке демонстрировался как чудо техники маятник с периодом в 30 минут, а тут — 30 часов! Какая у такого маятника должна быть приведенная длина! Но пока мы просто запомним эту цифру — 30 часов.

Известно, что среди звезд имеются так называемые цефеиды — звезды, периодически меняющие свою яркость. Периоды цефеид бывают очень различны, от нескольких часов до нескольких недель и даже месяцев. Что это — курьез? Нет, не курьез. Открытие цефеид — одна из самых замечательных страниц в науке. Удалось установить, что их период является однозначной функцией яркости. Тем самым открылись замечательные возможности. Пусть, например, мы наблюдаем две цефеиды с одинаковым периодом, скажем 10 дней, но яркость одной в 10 000 раз меньше, чем другой. Это значит, что она в 100 раз дальше от нас. Таким образом, цефеиды оказались как бы вехами, расставленными по Вселенной. Они позволили определять колоссальные расстояния — до сотен и тысяч световых лет, на которых параллаксы ничтожны: ведь до Солнца всего 8 минут. Тип колебаний цефеид еще не найден. Существуют теории Эддингтона, Джинса и других, но они объясняют только часть вопросов.

Совершенно те же методы, которые применяются для колебаний связанных систем, оказываются пригодными и в вопросах возмущений движения планет, например влияния Юпитера на движение Сатурна. Кому бы пришло в голову, что этот вопрос связан с вопросом о взаимодействии и колебаниях молекул твердого тела? Этот последний вопрос очень продвинут, но я все же не считаю его до конца решенным. При его развитии теория смогла объединить акустические и тепловые колебания. Я считаю это фундаментальным шагом, не меньшим, чем, скажем, электромагнитная теория света.

Я говорил до сих пор о двух кругах вопросов — о кинематике колебаний и о собственных колебаниях. Но есть третий и, быть может, наиболее типичный для всего учения о колебаниях. Это вопросы *резонанса*.

Насколько обширна эта область, вряд ли стоит говорить. Я мог бы повторить всем известные примеры — о мальчике, раскачивающем тяжелый колокол (пример, который идет от Галилея и повторен Гельмгольцем), о резонансе камертонов, где мы имеем сотни колебаний в секунду, о настройке радиоприемников, где речь идет уже о миллионах колебаний, о фраунгоферовых линиях, где частоты измеряются уже многими миллиардами.

Давно прошло то время, когда в технике делали только статические расчеты. Сколько было тогда неожиданностей и несчастий! Можно вспомнить, например, провал цепного моста в Испании в наполеоновские времена, провал египетского моста через Фонтан-

ку в Петербурге, когда по нему в ногу проходила гвардейская часть, разрывы валов машин из-за крутильных колебаний и много других. Интересно, что такого рода случаи встречались до самого последнего времени. В конце или в середине 30-х годов был спущен на воду самый быстрый по тому времени огромный французский пароход «Нормандия». На полном ходу он весь так сильно вибрировал, что через несколько рейсов его пришлось поставить в док. В 1900 г. на построенных тогда наших миноносцах «Баян» и «Громобой» оказалась такая вибрация, что мины срывались со своих мест, стрельба была невозможна. А. Н. Крылов первый указал на то, что дело было в резонансе, а раз известна причина, то можно и устранить все неприятности — изменить скорость, если же нельзя это сделать, то как-либо иначе.

Разумеется, резонанс вреден не всюду. Можно было бы привести колоссальное количество примеров его использования. Достаточно указать на радио, где усиление амплитуды при резонансе является одним из основных эффектов.

Но есть еще один момент, который иногда упускают из виду,— фазовые соотношения при резонансе, переход от колебаний в фазе с действующей силой к колебаниям в противофазе при изменении частоты силы от меньшей, чем собственная частота, к большей. С этим также связано много разнообразных явлений в самых различных областях.

Через греков до нас дошли записи вавилонских астрономов, в частности запись о соединении Сатурна и Юпитера с одной из ярких звезд. Когда в XVII в. Эйлер, Д'Аламбер и Клеро развили на основе ньютоновых законов теорию движения планет, то, естественно, хотели узнать, верна ли эта теория. Однако поди жди, пока осуществится какое-либо из ее предсказаний. Но можно произвести расчет в обратном направлении — от сегодняшнего дня на прошедшее. И вот соединение Сатурна и Юпитера, наблюденное вавилонянами, вычислили. Оказалось, что оно на 1° должно было происходить в другом месте, а это лежало за пределами ошибки наблюдений вавилонян. Положение было неясным, пока Лаплас в конце XVIII в. не обратил внимания на то, что периоды обращения Юпитера и Сатурна вокруг Солнца относятся почти точно, как 5 : 2. Поэтому здесь возникают своеобразные резонансные явления, и если их учесть, то получается полное согласие с наблюдениями вавилонян. Тем самым Ньютон полностью восторжествовал.

Фазовые соотношения при резонансе играют наряду с возрастанием амплитуды колоссальную роль и в житейских вопросах. Известно, что при качке корабль иногда зарывается носом, а иногда нет, хотя колебания одинаково сильны. Между тем оголение винтов чрезвычайно вредная вещь, вызывающая порчу машин. Первую

теорию корабля дал Эйлер, но выяснение только что указанного вопроса было дано лишь А. Н. Крыловым.

Играет ли здесь роль резонанс? Да, конечно. Но важна еще и фаза, т. е. важно то, больше или меньше собственная частота качаний корабля, чем частота набегающих на него волн. Вообще выяснение роли скорости хода и курса корабля, учет допплер-эффекта — все это было сделано А. Н. Крыловым.

Но все, что здесь имеет место при резонансе, можно слово за словом перевести на оптический язык в вопросе об аномальной дисперсии. Я остановлюсь, однако, на одном менее известном явлении, связанном с фазой при резонансе,— на лунных приливах и отливах.

Самое примитивное непонимание этого явления иллюстрируется рис. 1. Это, конечно, совершенно неверно: так было бы лишь в статическом случае, если бы все было неподвижно. В действительности — как учат в книжках — из-за ускорения получается выпуклость и с диаметрально противоположной стороны, как показано на рис. 2. Однако и эта картина не совсем верна, потому что в действительности выпуклости повернуты на некоторый угол и при наиболее высоком стоянии Луны еще не имеет места максимум прилива (рис. 3).

В чем же здесь дело? Вспомним, что океан — это маятник с периодом в 30 часов. А лунный период, т. е. полуоборот Луны вокруг Земли, составляет около 12 часов. Таким образом, период силы короче собственного, и мы имеем колебания со сдвигом фазы, близким

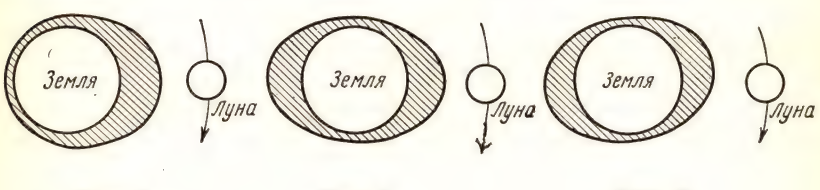


Рис. 1 Рис. 2 Рис. 3

к 180°. Только это и может дать единственно правильную динамическую теорию приливов.

Могу ли я считать, что сумел дать вам некоторое представление о громадном диапазоне теории колебаний? Можно было бы привести еще большое количество примеров, но уже и этих достаточно, чтобы увидеть разнообразие охватываемых явлений, в которые учение о колебаниях вносит единообразное понимание, чтобы увидеть универсальность этой дисциплины. Но если это так, то у вас естественно может возникнуть вопрос: все же не ошибаемся

ли мы, придавая теории колебаний такое большое значение? Ведь если бы дело было так, как я говорю, то она должна была бы играть первостепенную роль в развитии всей физики. Но оно так и есть.

Если вы посмотрите историю физики, то увидите, что главные открытия по существу были колебательными. Я постараюсь показать это теперь, хотя я и не хотел бы заходить так далеко, как английский математик и философ Уайтхед, который говорит, что рождение теоретической физики связано с применением к различным вопросам понятия периодичности.

Кстати, вы можете сказать: вот все время говорится о колебаниях и еще долго будет говориться, а не дано *определения*, что такое колебания. Я считаю, что давать определения — это трудная и неблагодарная задача. Вы дадите какое-то определение, а через два года его измените. Мне кажется, что важно и целесообразно другое. Важно указывать общие идеи, выделять *руководящие точки зрения*, которые позволяют нам *объединять целый класс проблем* наряду с *указанием тех фактических явлений*, о которых идет речь, с тем чтобы вызвать в памяти все их разнообразие. Так вот, одной из таких общих, руководящих точек зрения в учении о колебаниях является *периодичность*. Периодические явления — явления колебательные, но обратное неверно. Несомненно, то, что мы называем сегодня колебательными процессами, крайне разнообразно. Но развилось учение о колебаниях, по-видимому, на понятии — сегодня оно является частным понятием в области колебаний — *периодических* или приблизительно периодических (но не *почти*-периодических!) процессов [1]. Уже самое поверхностное наблюдение над явлениями природы привело нас к выделению таких процессов. Удары сердца и дыхание в нас самих, смена дня и ночи, смена времен года — все это периодические процессы. Ясно, что периодические, повторяющиеся явления легче изучать, чем однократные. Вообще я считаю, что явления, принципиально неповторимые, происходящие принципиально только один раз, не могут быть объектом изучения. Метеорология поэтому и находится в плохом состоянии, что в изучаемых ею явлениях нет периодичности. В теории приливов положение трудное, но все же лучше, так как некоторая периодичность существует.

Периодические явления начали входить в рассмотрение в XVI в. Коперник свел *сложные* видимые движения планет на их периодическое обращение вокруг Солнца. Затем Кеплер открывает свой знаменитый закон, который связал продолжительность этого обращения планет — *период* — с длиной большой полуоси их орбит,

76

---

[1] [Понятие почти-периодических процессов (функций) развилось позднее.]

закон, который лег в основу ньютоновского закона всемирного тяготения. Галилей наблюдает в Пизанском соборе качания паникадил и открывает изопериодичность малых колебаний. Он заложил учение о колебаниях маятника и сам кладет этот периодический процесс в основу измерения времени. Следует помнить, что часов ведь тогда не было, тогда не умели точно измерять время. Уже через два года после этих наблюдений Галилея Гюйгенс начал приспосабливать маятник к часам. Точные часы были необходимы для мореплавания, для определения долгот. В 1714 г. английское Адмиралтейство объявило премию в 20 тыс. фунтов стерлингов тому, кто построит хронометр с неточностью хода не более 3 секунд в сутки. Только через 45 лет Гаррисон построил такой хронометр (деньги ему, конечно, выдали не сразу, а еще через 10 лет). Важность этого дела нельзя преуменьшить: точные хронометры — это развитие мореплавания. Конечно, теперь их заменяет радио, но тогда это был жизненно важный вопрос.

Маятник сыграл большую роль и в создании общей теории относительности. Альфа и омега этой теории — равенство инертной и весомой масс. Это равенство впервые было установлено Ньютоном путем опытов с маятниками. С их помощью была исследована форма Земли. Измерения Рише (1672 г.), проведенные в Кайене и в Париже, показали, что для синхронности колебаний длина маятников должна отличаться в этих местах на 2,5 *мм*.

Еще бо́льшую роль сыграло учение о сложном маятнике, вопрос о нахождении центра качаний. На задаче о маятнике, как простом, так и особенно сложном, на задаче об определении периода такого маятника (вероятно, это одна из самых важных проблем в XVII и XVIII вв.) развился, по существу, вопрос о законе «живых сил» (Гюйгенс) и о сохранении энергии.

ВТОРАЯ ЛЕКЦИЯ

(28.III 1944 г.)

Прошлый раз я кратко и в общих чертах дал обозрение громадного материала, охватываемого теорией колебаний, а с другой стороны, тоже кратко начал исторический обзор того, какие этапы развития физики связаны с теорией колебаний. Сегодня я продолжу этот обзор.

Я указал, что Коперник вскрыл периодичность в обращении планет. Галилей и Гюйгенс изучали маятник. В этой проблеме по

существу впервые возник вопрос о сохранении энергии. Но в это же примерно время изучались и другие, я сказал бы, более скромные (с общей точки зрения) вопросы. В 1636 г. Мерсен формулировал все законы колебания струн, хотя и в непривычной для нас теперь форме. Он установил четыре положения: при одинаковых весах натягивающего груза числа колебаний (высоты тонов) относятся обратно длинам струн и т. д. По существу, он установил все, что содержится в формуле

$$n = \frac{1}{2l}\sqrt{\frac{T}{\rho}},$$

которую известный математик Тэйлор дал только в 1713 г. Эта формула сыграла большую роль.

Ньютон изучал распространение звука, т. е. периодических сгущений и разрежений воздуха, и открыл основной закон этого явления, чем было положено основание всему учению о распространении колебательных процессов в среде. Он получил формулу для скорости распространения

$$v = \sqrt{\frac{dp}{d\rho}},$$

но принял изотермический закон $p = c\rho$ вместо адиабатического $p = c\rho^{\gamma}$, что в дальнейшем сделал Лаплас.

Именно Ньютон, а не Гюйгенс — это недостаточно известно — первый делает замечательное открытие, что в основе монохроматического света лежит *периодический* процесс. На это открытие опирается, по существу, дальнейшее развитие всей оптики.

Надо было объяснить открытые самим Ньютоном кольца. Ньютон сказал, что они возможны только потому, что в монохроматическом свете имеется периодичность, и определил период — длину волны λ. Вместе с тем он считал, что свет — это поток частиц, и вот тут он и выдвинул свою знаменитую гипотезу «приступов» — периодических «припадков» у световых частиц легкого прохождения и легкого отражения. На этой основе, видимо, было трудно провести последовательную теорию, однако самое открытие периодичности принадлежит Ньютону.

Следующий этап связан с именами Юнга и Френеля. Ньютон умер в 1727 г., но авторитет его был так велик, что эмиссионная теория держалась почти сто лет. В 1801 г. в своем бакерианском чтении Юнг сказал: свет — это колебания очень тонкой среды — эфира, заполняющего всю Вселенную. Колебательное движение возбуждается в этом эфире всякий раз, когда тело начинает светиться. Говоря, что свет — это колебания, он примкнул к Ньютону, вводя эфир, он примкнул к Гюйгенсу. Частота колебаний, по

Юнгу, определяет цвет света. Он дал принцип известного объяснения интерференции и заложил основы волновой оптики, однако признана она была далеко не сразу. Господствовавшие в то время умы — Лаплас, Пуассон, Био — не хотели верить этой теории и решили ее убить, поставив конкурсную тему — объяснение дифракции света. У Ньютона дифракция объяснялась просто — протяжением частиц к краю экрана.

Ответом на это был мемуар Френеля, представленный в 1819 г., в котором он не только объяснил, но и количественно рассчитал дифракцию. Это была не смерть, а победа волновой теории.

Френель объяснял все явления, пользуясь зонами. На заседании комиссии Пуассон возразил, что, по Френелю, в центре тени круглого экрана должно получаться светлое пятно. Непосредственно после заседания Араго показал на опыте, что это так и есть. С тех пор волновая теория окончательно восторжествовала. В 1850 г. Физо непосредственно измерил скорость света и, как думали тогда, доказал волновую теорию. Таким образом, колебания и волны завоевали громадную область физики — оптику.

Однако и после Френеля было еще немало трудностей и противоречий в вопросе о свойствах упругого эфира. В силу поперечности световых колебаний он должен был обладать свойствами твердого тела. С другой стороны, он не оказывает никакого сопротивления движениям планет. В другой форме эти противоречия занимали уже Ньютона. Их преодоление пришло из совсем другого круга идей, связанного с именами Фарадея и Максвелла.

19 октября 1861 г. Максвелл пишет свое знаменитое письмо Фарадею об электромагнитной теории света, которая и устранила все трудности механических теорий. С Максвеллом было обратное тому, что рассказывает легенда о Ньютоне. Ньютон будто бы 20 лет держал в секрете закон тяготения, так как значения ускорения для Луны и для падающего камня не сходились; он ждал, пока не будет более точно измерен радиус Земли. Максвелл сразу же заключил и высказал, что свет есть электромагнитные волны, опираясь на следующее поразительное совпадение. В его теории есть величина c — отношение электростатической единицы заряда к электромагнитной, равная 193,088 *миль/сек.* С такой скоростью, по теории Максвелла, должны распространяться электромагнитные волны. По прямым измерениям Физо, цитирует Максвелл, скорость света равна 193,188 *миль/сек* (1 *миля* = 1,6 *км*). Любопытно то, что оба эти числа неверны на 6000 *миль*, но случайно ошибки в обоих в одну сторону! В последующей публикации Максвелл дает уже другие числа, хуже совпадающие. К счастью, однако, это расхождение не разочаровало Максвелла.

Фарадей и Максвелл создали эпоху в физике. Но на чем укрепилась их теория? Опять-таки на колебаниях. Для развития фара-

дей-максвелловских представлений об электрических и магнитных явлениях, для создания электромагнитной теории света решающее значение имело теоретическое изучение колебательных периодических процессов самим Максвеллом и осуществление электромагнитных колебаний Герцем. Основные опыты были сделаны Герцем в 1886 г. Герц написал о своих опытах письмо Гельмгольцу. Ответ был краток: «Манускрипт получил. Браво!» Эта оценка кое-что означала.

Любопытно, что в дневнике Герца имеются (в двух местах) упоминания о распространении волн в трубах. Так, например, 17 мая 1888 г. он пишет: «Начал работать по вопросу о распространении в металлических трубах».

После первых увлечений максвелловской теорией начались неполадки: надо было объяснить дисперсию. Теория Максвелла не могла этого сделать. Она занималась распространением электромагнитных волн, но не рассматривала их источник. Здесь на сцену появляется Лоренц.

Лоренц предположил, что источником света является упруго связанный электрон, т. е. электрон-осциллятор, сидящий в атоме и обладающий собственной частотой. Таким образом, колебания были внесены в атом и это позволило объяснить дисперсию.

Сейчас мы стоим на квантовой точке зрения, но те же колебательные явления, как мне кажется, заставили бросить все старое и перейти к квантам. Спектральные серии не удавалось объяснить классически. При тепловом равновесии каждая степень свободы имеет определенную энергию. Спектральная серия содержит бесконечно много линий, каждая линия — от своего осциллятора, а следовательно, из-за огромного числа степеней свободы должна быть колоссальная теплоемкость. Опыт не давал ничего подобного. Эти соображения, вносящие ясность в вопрос, высказал Релей. Он почерпнул их из серий и equipartition [закона равномерного распределения энергии по степеням свободы], т. е. опять-таки из колебательных вещей. Кельвин хотел показать, что equipartition — неверный закон, и придумывал для этого разные примеры. Релей показывал, что неверны примеры и что так просто не отделаешься, а нужен радикальный пересмотр всего. Однако теорию квантов он встретил без энтузиазма. В октябре 1911 г. он писал: «Но я признаюсь, мне не нравится это решение загадки. Конечно, я не могу ничего возразить против того, чтобы проследили следствия теории квантов энергии, которая уже дала в руках способных людей ряд интересных результатов. Но я затрудняюсь принять это в качестве картины того, что происходит в действительности».

И в разрушении классических представлений, и в создании волновой механики колебания играли основную роль. Каждое из названных открытий, относящихся именно к периодическим явле-

ниям, как ни различны они сами по себе, составило эпоху в развитии физики. Недаром вопросы колебаний имеют такое большое значение. Каждый успех в этой области, несомненно, имеет гораздо более широкое значение, чем те специальные вопросы, на которых он был достигнут, и не удивительно, что теория колебаний выделена теперь в самостоятельную дисциплину.

Я хотел бы отметить еще одну черту теории колебаний. В ней мы имеем, пожалуй, наиболее сильное взаимодействие с математикой. До последних 30—40 лет весь аппарат был *линейным*, соответственно и колебания были *линейные* колебания. Дифференциальные уравнения в частных производных и краевые задачи, интегральные уравнения, разложение произвольных функций в ряды по ортогональным функциям и, в частности, ряды Фурье,— в развитии всех этих математических средств физические запросы играли не последнюю роль. Но только благодаря развитию этих математических дисциплин сделалось возможным углубленное понимание основных физических колебательных процессов.

Однако огромное количество явлений принципиально не может быть хорошо рассмотрено с помощью линейных уравнений. Вопросы радио, музыкальные инструменты (кроме ударных и щипковых), автоматика и регулирующие механизмы, сосуществование видов в биологии, вопросы трения (например, колебания, возникающие при сухом трении, давшие повод Боудену и Лебену в 1939 г. утверждать, что само трение имеет скачкообразный характер, хотя еще раньше С. Э. Хайкин показал, что дело вовсе не в этом), динамика полета, поведение цефеид — все это вопросы, не поддающиеся линейной трактовке.

Нелинейный математический аппарат был известен ранее. Он был создан Пуанкаре и Ляпуновым, но для других задач. Исследования Ляпунова вообще были математическими. В конце своего труда он писал, что он должен был бы привести примеры, но действительных примеров нет, а он не хочет искусственно их создавать. Таким образом, математика предвосхитила здесь потребности теории колебаний. В настоящее время несомненно, что задача и физиков, и математиков состоит в том, чтобы разрабатывать нелинейные методы дальше.

На этом я закончу исторический обзор.

Как я уже говорил, я хотел бы взять лишь несколько основных вопросов и на них показать, в чем заключаются трудности, в чем состоит центральный пункт вопроса и как результаты его решения применяются в разных областях. Углубляться, делать все выводы я, конечно, не могу.

Я сказал, кроме того, что точное определение той или иной дисциплины — очень трудное и нецелесообразное дело. Примером может служить разграничение физики и химии, попытки поставить

между ними колючую проволоку. Важно другое, важны руководящие точки зрения, объединяющие ряд проблем, позволяющие создать цельную стройную концепцию, связать в одно целое разнородные явления и отсюда получить нить для дальнейших исследований. Указание и выяснение таких точек зрения — это хорошее дело. Но можно ли их найти в теории колебаний? Да, можно.

Чтобы понять колебательные процессы, надо знать полностью их протекание во времени. В этом их отличительная черта. Если я хочу узнать работу, производимую быстро движущимся телом, то я не интересуюсь историей процесса. Мне достаточно знать скорость тела в данный момент, и я тогда вычисляю $A = \frac{mv^2}{2}$. В колебательных процессах так делать нельзя. Если нас интересует действие силы на резонатор, то надо знать *всю форму* колебания силы. Два колебательных процесса могут различаться по своей форме лишь в узком интервале времени, и все же это уже *разные* процессы.

Периодические процессы заданы тем, что

$$f(t + T) = f(t)$$

для *всех* t; T — период. Это определение имеет смысл, только если оно выполняется для любого t от $-\infty$ до $+\infty$. Существует ряд типичных, наиболее удобных, *стандартных* форм периодических процессов. Среди них очень важная синусоидальная форма

$$y = A \cos(\omega t + \varphi).$$

Все равно, что̀ такое y. Важна сама форма (простое, гармоническое, монохроматическое колебание). Для выкладок иногда удобнее комплексная форма

$$y = Ae^{i\omega t},$$

где A — комплексная величина. При этом подразумевается, что потом, т. е. в результатах, мы перейдем к действительной части.

Почему это одна из важных стандартных функций? Сразу ответить на это трудно, но мало-помалу мы в этом убедимся. Прежде всего можно указать на то, что в природе очень много явлений, протекающих по этому закону,— колебания маятника без затухания, простой тон, монохроматический свет, правильная зыбь на поверхности воды, колебания в радиопередатчике. Второе обстоятельство, на которое следует указать, состоит в том, что функции, описывающие процессы иной формы, можно представить рядом или интегралом Фурье, т. е. разложить их на синусообразные составляющие. Но тут мы и наталкиваемся на вопрос, на который обычно

мало обращают внимания, а из этого проистекает ряд плохих последствий.

Я сказал, что синусоида важна нам потому, что она часто встречается в действительности. Но ведь синусоида определена тем, что A, ω и φ постоянны во все времена t, т. е. от $-\infty$ до $+\infty$. Таким образом, маятник колеблется по синусоиде *только* в том случае, если он колеблется по указанному закону от $-\infty$ до $+\infty$. То же относится и к простому тону, и к монохроматическому свету, и т. д. Ясно, что в природе таких явлений нет, всегда есть начало и конец колебания, т.е. синусоида *никогда* не встречается. Более того, часто самое применение колебаний требует изменения A, ω или φ; так обстоит дело при измерении скорости света, так обстоит дело в музыке, при телеграфировании и т. д. Вместо синусоиды мы имеем тогда отображающую функцию вида $A(t)\cos(\omega t + \varphi)$ или же функцию, в которой ω или φ зависят от t. Но говорить о синусоиде с переменной амплитудой (частотой, фазой) — это значит вступать в противоречие с определением. Ведь синусоида определена *однозначно* только при условии постоянства A, ω и φ при *всяком* t от $-\infty$ до $+\infty$.

Спрашивается, что же делать: мы всегда и успешно пользуемся синусоидой, а в природе ее никогда нет? Это серьезный вопрос, и подойти к нему надо также серьезно. Его можно формулировать так: при каких условиях, несмотря на то, что синусоидальных процессов в природе не бывает, можно с достаточной точностью пользоваться косинусом и синусом?

Можно, конечно, сказать, что при *достаточно медленном* изменении A, ω или φ, как это имеет место, например, в тех случаях, когда они меняются «по долгу службы» (пение, радиотелефония), я могу решать вопросы так, как если бы они были постоянны. Но тогда тотчас же возникает другой вопрос: что значит «достаточно медленно»? A priori почти никогда нельзя сказать, достаточно ли медленно меняются A, ω, φ, и это *типично* для колебательных процессов. Необходимо знать, *на что это колебание действует*: при одной и той же $A(t)$ для одного приемника изменение может оказаться достаточно медленным, для другого — нет.

Вместе с тем *благодаря* изменению A, ω, φ возникают *новые* вопросы, например вопрос о *распространении* таких колебаний, о соответствующих волнах, акустических и ультраакустических, оптических, радио, волнах на поверхности воды, волнах в каналах, упругих волнах (например, землетрясения) и т. д.

Начнем с простого и конкретного вопроса. Мы нарочно накладываем на волну какой-то признак, скажем меняем амплитуду для того, чтобы дать сигнал. Мы имеем, таким образом, синусоиду с медленно меняющейся амплитудой, и вопрос заключается в том, как будет распространяться этот сигнал или *группа* волн (такую

волну называют группой). Ведь нам известно только, как распространяется синусообразная волна.

Но здесь приходит на помощь то обстоятельство, что главные типы волн подчиняются *линейным*, *однородным* дифференциальным уравнениям с *постоянными* коэффициентами. Благодаря этому, как мы увидим далее, синусообразные волны и оказываются особо выделенными.

Рассмотрим конкретные примеры, ограничиваясь одномерным случаем. Однородная «мягкая» струна при постоянном натяжении. Ее поперечное отклонение подчиняется уравнению

$$\frac{\partial^2 y}{\partial t^2} - a^2 \frac{\partial^2 y}{\partial x^2} = 0.$$

Поперечные (изгибные) колебания тонкого «жесткого» стержня описываются таким уравнением:

$$\frac{\partial^2 y}{\partial t^2} - a^2 \frac{\partial^4 y}{\partial x^4} = 0.$$

Оптические (электромагнитные) колебания в однородной среде — поляризованная волна. Пусть E и H — электрический и магнитный векторы, а s — среднее отклонение электрона. Тогда уравнения, которым подчиняются эти величины, будут

$$\frac{1}{c}\frac{\partial H}{\partial t} = -\frac{\partial E}{\partial x}, \qquad \frac{1}{c}\left(\frac{\partial E}{\partial t} + Ne\dot{s}\right) = -\frac{\partial H}{\partial x},$$

$$m\ddot{s} + fs = eE,$$

где e и m — заряд и масса электрона, N — число электронов в 1 $см^3$, fs — квазиупругая сила, возвращающая колеблющийся электрон в положение равновесия.

Во всех трех случаях мы имеем уравнения с постоянными коэффициентами. Такие уравнения обладают замечательным свойством: для каждой из величин они имеют частное решение вида

$$y = e^{i(kx - \omega t)},$$

где ω и k — постоянные, но не любые, а связанные между собой, так что $\omega = \varphi(k)$, причем $\varphi(k)$ определяется видом уравнения. Подставьте это частное решение, и получатся алгебраические уравнения, определяющие связь между ω и k.

Существенно еще одно обстоятельство: сумма частных решений благодаря линейности уравнений также есть частное решение. Это *принцип суперпозиции*, играющий чрезвычайно важную роль в теории колебаний, хотя почти всегда его применяют неправильно (но об этом речь пойдет дальше).

К сожалению, есть много случаев в физике, где дело обстоит не так просто. Например, волны при больших отклонениях от положения равновесия описываются *нелинейными* уравнениями, волны в неоднородной среде — уравнениями с *непостоянными* коэффициентами и т. п. К таким случаям наша теория не относится.

Пусть при $t = 0$ задано начальное распределение $y = e^{ikx}$. Тогда при любом t получим $y = e^{i(kx-\omega t)}$, т. е. волну, распространяющуюся со скоростью

$$v = \frac{\omega}{k}.$$

Это ясно. Легко показать, что в отсутствие поглощения, т. е. при незатухающих волнах, если $\omega = \varphi(k)$ удовлетворяет нашим уравнениям, то $-\omega$ тоже удовлетворяет. Мы получаем, другими словами, на выбор два направления распространения вдоль оси x [причем то или иное из них определяется заданием при $t = 0$ не только y, но и $\dot{y}$]. Ограничимся далее одним направлением — в сторону возрастания x.

Нас интересует, однако, другой вопрос: как будет распространяться волна, заданная при $t = 0$ распределением y:

$$y = f(x)\, e^{ik_0x}?$$

Прежде всего заметим, что *всегда*, *во всех* вопросах распространения мы сталкиваемся именно с *этой* задачей, ибо никогда $f(x)$ не бывает константой. Правда, $f(x)$ обычно меняется медленно, гораздо медленнее, чем e^{ik_0x}, но все же это не постоянная величина.

До Стокса говорили так: $f(x)$ — почти постоянна, а значит, волна будет распространяться так, как будто $f(x)$ — постоянная величина

$$y = f(x)\, e^{i(k_0x-\omega_0t)},$$

т. е. скорость распространения такая же, как у синусоидальной волны

$$v = \frac{\omega_0}{k_0}.$$

Но это как раз и неверно. В среде с дисперсией, *как бы медленно ни изменялась* $f(x)$, волна так не распространяется, и впервые на это указал Стокс.

Наши основные уравнения имеют частное решение в виде бесконечной синусоидальной волны. Как распространяется синусоидальная волна, мы знаем. На наше счастье, существует теорема Фурье, и поэтому мы можем решить задачу так: представим $f(x)$

в виде интеграла Фурье

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k)\, e^{ikx} dx,$$

где

$$g(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)\, e^{-ikx} dx.$$

Тогда начальное отклонение можно записать в виде

$$y(x, 0) = e^{ik_0x} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k)\, e^{ikx} dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k)\, e^{i(k+k_0)x} dk.$$

Каждый элемент интеграла $g(k)\, e^{i(k+k_0)x}\, dk$ представляет синусоидальную волну, протяженную по x от $-\infty$ до $+\infty$, волну с *постоянной* амплитудой. Мы имеем непрерывный набор таких синусоидальных функций, и к каждой из них уже может быть применена наша теория.

Таким образом, при $t > 0$ мы получим

$$y(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k)\, e^{i[(k+k_0)x-\omega t]} dk, \tag{1}$$

где ω является определенной функцией от $k + k_0$.

Я нарочно сделал все подробно, потому что часто эти выкладки обходят и говорят о них лишь на словах. В следующий раз мы продискутируем эту формулу, из которой без дальнейшего следует все учение о групповой скорости.

Замечу только, что теперь уже не представляет труда дать определение того, что значит, что $f(x)$ медленно изменяется. Это значит, что $g(k) \neq 0$ лишь при малых значениях k, т. е. в разложении $f(x)$ будут играть роль только значения $k \ll k_0$. Малость этих значений k по сравнению с k_0 и дает меру медленности $f(x)$. Если условие $k \ll k_0$ выполнено, то вся дискуссия может быть доведена до конца.

ТРЕТЬЯ ЛЕКЦИЯ

(7.IV 1944 г.)

Мы занялись вопросом о распространении группы волн, на который еще в 1876 г. обратил внимание Стокс [1]. Пользуясь тем выражением для группы волн, которое мы получили в прошлый раз, определим величину групповой скорости.

Напомню, что при соответствующих начальных смещениях и скоростях в момент $t = 0$ возмущение, имеющее вид

$$y(x, 0) = f(x)\, e^{ik_0x},$$

распространяется затем в виде группы волн, которую можно записать формулой (1).

Как я уже говорил, медленность изменения $f(x)$ («амплитуда» начальной волны), ведет к тому, что $g(k)$ отлично от нуля только при небольших k. Поэтому ω, которая является функцией $k + k_0$, можно приближенно представить в таком виде: $\omega = \omega_0 + \left(\frac{\partial\omega}{\partial k}\right)_0 k + \ldots$, и ограничиться в этом разложении членом первого порядка относительно k.

Тогда наше выражение для группы волн можно переписать так

$$y(x, t) = e^{i(k_0x - \omega_0 t)} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k)\, e^{ik\left[x - \left(\frac{\partial\omega}{\partial k}\right)_0 t\right]} dk,$$

откуда видно, что

$$y(x, t) = f\left[x - \left(\frac{\partial\omega}{\partial k}\right)_0 t\right] e^{i(k_0x - \omega_0 t)},$$

где f — медленная функция, которая задает начальную форму сигнала, но которая теперь зависит не от x, а от двучлена, стоящего в скобках. Таким образом, множитель при $e^{i(k_0x - \omega_0 t)}$ можно рассматривать как «переменную амплитуду» волны, распространяющейся с

---

[1] [В конспекте Л. И. Мандельштама после упоминания о Стоксе стоит фраза: «Гамильтон в 1839 г. сделал сообщение в Ирландской академии, обещал напечатать подробнее, но доклад не был опубликован». В 1839 г. в Proc. Roy. Irish. Acad., I, 267 и 341 появились две заметки Гамильтона, в которых для обычной в то время дискретной модели эфира он ввел не только «скорость распространения колебательного движения», но и скорость «первого возмущения», т. е. на современном языке — групповую скорость и скорость предвестника. Более подробное рассмотрение этих вопросов содержится в рабочих тетрадях Гамильтона, опубликованных в 1940 г. (The mathematical papers of sir William Rowan Hamilton, V. II, Dynamics. Cambridge University Press, 1940).]

фазовой скоростью $v_0 = \frac{\omega_0}{k_0}$. Сама эта «амплитуда» распространяется, однако, с другой скоростью, а именно:

$$u = \left(\frac{\partial\omega}{\partial k}\right)_{k=k_0}.$$

Эту величину и называют *групповой скоростью.*

Проделанный нами вывод показывает, что понятие групповой скорости можно ввести только для определенного вида возмущения, обладающего достаточно узким спектром синусоидальных бесконечных волн ($k \ll k_0$), и только при наличии определенных свойств у рассматриваемой однородной среды. Эти свойства должны быть таковы, чтобы для тех длин волн, которые составляют группу, с достаточным приближением можно было рассматривать частоту колебаний ω как *линейную* функцию волнового числа k. Грубо говоря, все рассмотрение возможно и применимо лишь в том случае, если «амплитуда» очень мало меняется на длине основной волны $\lambda_0 = \frac{2\pi}{k_0}$ и, кроме того, показатель преломления среды не испытывает для волн вблизи λ_0 резких изменений. Это, конечно, только слова, но формулы и их вывод, если их понимать правильно, говорят о том же гораздо точнее.

Действительно, напишем следующий член разложения ω по степеням k. Это будет $\left(\frac{\partial^2\omega}{\partial k^2}\right)_0 \frac{k^2}{2}$. В показателе экспоненты под знаком интеграла этот член даст дополнительную фразу $\left(\frac{\partial^2\omega}{\partial k^2}\right)_0 \frac{k^2 t}{2}$, и для того, чтобы можно было ее не учитывать, необходимо потребовать

$$\left|\left(\frac{\partial^2\omega}{\partial k^2}\right)_0 \frac{k^2 t}{2}\right| \ll 1.$$

Нетрудно переписать это условие несколько иначе. Пусть наш сигнал («переменная амплитуда») пробежал путь X, потратив на это время $t = \frac{X}{u}$. Учитывая, что $\left(\frac{\partial^2\omega}{\partial k^2}\right)_0 = \left(\frac{\partial u}{\partial k}\right)_0$ и заменяя k через $\frac{2\pi}{\lambda}$, получаем

$$\left|\frac{\pi}{u}\left(\frac{\partial u}{\partial\lambda}\right)_{\lambda_0} X\right| \ll 1,$$

т. е. при наличии изменения групповой скорости по спектру расстояние X, до которого пригодно все наше рассмотрение, ограничено этим условием. Чем меньше $\left(\frac{\partial u}{\partial\lambda}\right)_{\lambda_0}$, тем больше допустимое X.

Прежде чем перейти к конкретным примерам, заметим еще следующее. Если $\omega = \alpha k$, где α — константа, то в среде нет дисперсии:

$$v = \frac{\omega}{k} = \alpha.$$

В этом случае

$$u = \frac{\partial\omega}{\partial k} = \alpha = v,$$

т. е. групповая скорость равна фазовой. Группа будет распространяться при этом так же, как и отдельная синусоидальная волна. С таким положением вещей мы сталкиваемся в случае звуковых волн в воздухе.

Если заменить ω через vk и вместо k ввести длину волны по формуле $k = \frac{2\pi}{\lambda}$, то для групповой скорости получится известная формула Релея

$$u = v - \frac{\partial v}{\partial \lambda}\lambda.$$

Отсюда, между прочим, сразу же видно, что групповая скорость

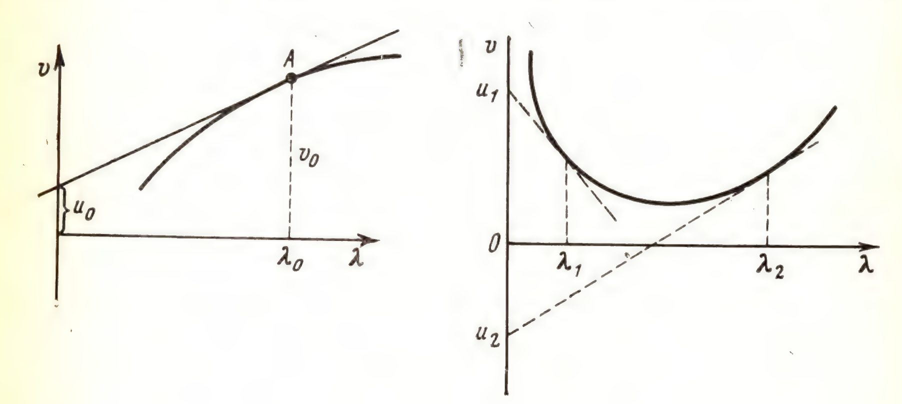


Рис. 4 Рис. 5

может быть как больше, так и меньше фазовой, в зависимости от знака $\frac{\partial v}{\partial \lambda}$.

Эренфест указал на простой графический способ нахождения u по кривой $v(\lambda)$. Как ясно из рис. 4, отрезок, отсекаемый на оси ординат касательной к кривой $v(\lambda)$, проведенной в какой-либо

точке A этой кривой, как раз равен $v_0 - \left(\frac{\partial v}{\partial \lambda}\right)_0 \lambda_0$, т. е. значению групповой скорости u_0 в точке A. Следовательно, если кривая $v(\lambda)$ достаточно круто поднимается, то групповая скорость может и обратиться в нуль, и сделаться отрицательной, как, например, в точке λ_2 на рис. 5. При отрицательной групповой скорости «медленно меняющаяся амплитуда» движется в сторону, противоположную направлению распространения волн, составляющих группу.

Как представить себе это наглядно? Флеминг приводит следующий пример: по плывущей барже вереницей бегут мальчики, спрыгивают с носа в воду и взбираются затем на корму. Если мальчики будут бежать в сторону, противоположную движению баржи, то мы и получим картину отрицательной групповой скорости (скорость баржи). Примером нулевой групповой скорости может служить эскалатор: ступени (играющие роль волн) движутся, а группа в целом, т. е. весь эскалатор, стоит на месте.

Перейдем теперь к реальным примерам.

Групповая скорость волн на поверхности воды в глубоком бассейне может быть получена из следующей известной зависимости между фазовой скоростью и длиной волны:

$$v^2 = \frac{g}{k} + k\frac{T}{\rho},$$

где g — ускорение силы тяжести, T — капиллярная постоянная жидкости, ρ — ее плотность. Можно взять два крайних случая:

1) гравитационные (длинные) волны, при которых k малы и $v^2 \approx \frac{g}{k}$. Тогда групповая скорость будет

$$u = \frac{v}{2};$$

2) капиллярные (короткие) волны, при которых k настолько велики, что $v^2 \approx \frac{kT}{\rho}$, и тогда

$$u = \frac{3v}{2}.$$

Из общей формулы для v^2 следует, что существует минимальная скорость поверхностных волн. Она соответствует длине волны $\lambda = 2\pi\sqrt{\frac{T}{\rho g}}$ и равна $v_{\min} = \sqrt[4]{\frac{4gT}{\rho}}$. Для воды это волна длиной около 1,7 *см*, и ее фазовая скорость равна примерно 23 *см/сек*. Наблюдались гравитационные волны длиной около $^3/_4$ *км* (период равен 23 *сек*), обладающие фазовой скоростью порядка 120 *км* в час.

В случае поперечных колебаний стержня мы имеем $v = ak$, и, следовательно,

$$u = 2v.$$

Рассмотрим теперь распространение электромагнитных (световых) волн в прозрачной среде. Мы пренебрегаем и поглощением, и поправкой Лоренца на взаимодействие атомных резонаторов (упруго связанных электронов). Тогда справедливы уравнения, которые мы писали в прошлый раз, и из них нетрудно получить следующее выражение для фазовой скорости:

$$v = \frac{c}{\sqrt{1 + \frac{a}{\omega_0^2 - \omega^2}}},$$

где $a = \frac{4\pi Ne^2}{m}$ и $\omega_0 = \sqrt{\frac{f}{m}}$ — собственная частота резонаторов.

Отсюда следует, что групповая скорость есть

$$u = c\,\frac{\sqrt{1 + \frac{a}{\omega_0^2 - \omega^2}}}{1 + \frac{a\omega_0^2}{\omega_0^2 - \omega^2}}.$$

Из этих формул видно, что при $\omega > \omega_0$ фазовая скорость больше скорости света в пустоте ($v > c$), а групповая — меньше ($u < c$). В том, что $v > c$, конечно, нет противоречия с теорией относительности, ибо фазовая скорость не есть скорость сигнала: бесконечная синусоидальная волна не может «передать» управляемого воздействия «отсюда — туда», а теория относительности требует скорости, меньшей c, только для таких воздействий.

Если электроны в среде свободны ($f = 0$, $\omega_0 = 0$), как это имеет место, например, в ионосфере, то из только что написанных формул находим:

$$v = \frac{c}{\sqrt{1 - \frac{a}{\omega^2}}} \quad \text{и} \quad u = c\sqrt{1 - \frac{a}{\omega^2}}.$$

Таким образом, здесь всегда $v > c$ и $u < c$, причем

$$uv = c^2.$$

Рассмотрим более внимательно законы распространения сигнала в среде.

Скорость — понятие, возникшее при описании движения частицы. Оно является совершенно ясным и имеет смысл при том условии, что существует возможность отождествления частицы, т. е. в любой точке пространства мы можем утверждать, что это та же самая частица.

При распространении волны мы имеем дело с перемещением не частицы, а *состояния*. Чтобы говорить о скорости, нужно иметь возможность и средства для отождествления *состояния*. В среде без дисперсии, когда фазовая скорость не зависит от длины волны, всякое возмущение распространяется без изменения формы, поэтому возможность отождествления здесь очевидна. Но в среде с дисперсией возмущение по мере распространения деформируется, и здесь уже нельзя без дальнейшего сказать, чему равна скорость. Нужно сначала определить, что мы в каждом таком случае будем называть скоростью распространения. Например, для движения облака нет однозначного понятия скорости. Это может быть и скорость края облака, и скорость его центра масс, и т. д. Необходимо каждый раз условиться, чтó мы считаем скоростью облака. Примерно так же обстоит дело и со скоростью возмущения.

То, что было найдено нами, относится к скорости распространения «переменной амплитуды», и эта скорость (групповая) имеет смысл только при условии неизменности группы при ее перемещении. В диспергирующей среде такой неизменности нет, но при выполнении условий, о которых мы говорили (достаточно медленное изменение амплитуды, т. е. малое ее изменение на длине волны, и достаточно пологий ход дисперсии), деформация группы также происходит медленно, и тогда для не слишком больших расстояний понятие групповой скорости *приближенно* описывает распространение группы.

Во всяком случае, всегда, когда есть дисперсия, понятие скорости теряет однозначность. Можно по-разному определить скорость, и одно из определений, являющееся адекватным при указанных ограничениях, — это групповая скорость. В сущности говоря, во всех опытах, в которых речь идет о непосредственном измерении скорости света, измеряется именно групповая скорость. В опыте Физо, где свет прерывается, давая цуги, это очевидно. Но так же обстоит дело и в опытах Фуко — Майкельсона, в которых скорость света определяется при помощи вращающегося зеркала. Луч света падает на вращающееся зеркало A, отражается от неподвижного зеркала B и, идя обратно, встречает зеркало A в другом положении, в силу чего возвращается под углом к первоначальному направлению. Кроме того, пучки, падающие на приближающиеся и удаляющиеся части зеркала, претерпевают изменение частоты при отражении (допплер-эффект). Изменение частоты и угол поворота являются мерой скорости света. Очевидно, что в прибор на-

блюдателя попадает волна, модулированная по амплитуде и частоте, а значит, и в этом случае также измеряется групповая скорость.

Существует еще один способ измерения скорости света, основанный на определении угла аберрации (Брэдлей). Наблюдая звезду в телескоп, необходимо наклонить его на угол $\varphi = v/c$ (рис. 7) от действительного направления на звезду (c — скорость света, v — составляющая скорости Земли, перпендикулярная к линии наблюдения). Что здесь измеряется, групповая или фазовая скорость?

На первый взгляд может показаться, что синусоидальная волна, приходящая от звезды, в этом опыте никак не видоизменена и что скорость света c, получаемая из измерения угла, является поэтому фазовой скоростью. Однако Эренфест указал на то, что и здесь есть нечто, что накладывает отпечаток на волну. Схематически распространение света через движущееся вдоль фронта волны отверстие объектива можно представить себе следующим образом. Свет проходит последовательно через две диафрагмы, и для того, чтобы увидеть через них звезду, над наклонить линию центров обеих диафрагм по отношению к направлению на звезду на угол φ (аберрация). Первая диафрагма выделяет *луч*, т. е. ее края как бы разрезают волну и в прибор попадает не непрерывная волна, а

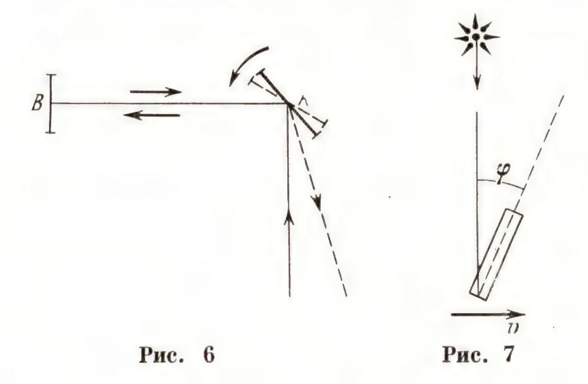


Рис. 6 Рис. 7

следующие друг за другом цуги волн (рис. 8). Такое ограничение пучка приводит к тому, что и в этом опыте φ определяется групповой скоростью. Конечно, рассуждение нужно подтвердить расчетом. Такой расчет был сделан и подтвердил, что и в этом случае измеряется групповая скорость [1].

В сущности говоря, фазовую скорость мы измеряем только из показателя преломления.

[1] [См. работу М. П. Свешниковой (ЖРФХО, 59, 377, 1927).]

Понятие групповой скорости существенно во многих задачах, касающихся распространения волн. Мы сталкиваемся с ним при рассмотрении допплер-эффекта в диспергирующей среде — вопрос, рассмотренный И. М. Франком; в вопросе о распространении электромагнитных волн в (пупинизированных) кабелях и трубах; в задаче о распространении волн на поверхности жидкости и сейсмических волн в земной коре.

Можно ли говорить о скорости импульса в диспергирующей среде? Если импульс очень короткий, т. е. обладает широким частотным спектром, то по мере распространения форма его будет сильно искажаться. Понятие групповой скорости будет в этом случае неопределенным. Конечно, в частных случаях при соблюдении известной осторожности можно ввести понятие скорости импульса, но никоим образом нельзя распространять это понятие на все случаи без разбора.

Возможно, однако, и другое определение понятия скорости. Так, например, в случае импульса мы можем говорить о скорости его *фронта*, т. е. о скорости переднего края. Нетрудно сообразить, что в любой диспергирующей среде скорость фронта всегда равна скорости света в вакууме. Ведь только наличие резонаторов, из которых построена среда, приводит к тому, что скорость волны меняется. Когда же речь идет о фронте импульса, то во время его прохождения резонаторы еще не успевают сдвинуться с места и поэтому их излучение не скажется на скорости фронта. Фронт любого сигнала проходит через среду, как через пустое пространство, и действие среды сказывается лишь *позади* фронта. Отсюда понятно, что скорость сигнала, как бы мы ее ни определяли, всегда меньше (или равна) скорости фронта, т. е. скорости света в вакууме, в согласии с требованием теории относительности.

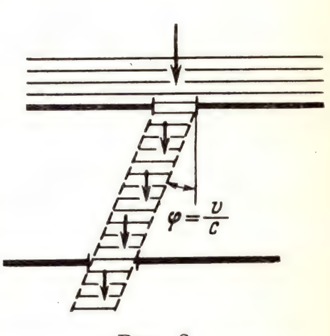


Рис. 8

О том, что получается, если *не* принимать это различие *понятий* скорости в расчет, видно из двух следующих разительных примеров.

Согласно лоренцову преобразованию, отсчеты времени t и t' в двух инерциальных системах координат, относительная скорость которых есть V, связаны так:

$$t' = \frac{t - \frac{V}{c^2}x}{\sqrt{1 - \frac{V^2}{c^2}}}.$$

Если в нештрихованной системе сигнал в момент t вышел из точки x и в момент $t = \Delta t$ пришел в точку $x + \Delta x$, то по часам штрихованной системы между выходом и приходом сигнала прошел промежуток времени

$$\Delta t' = \frac{\Delta t - \frac{V}{c^2}\Delta x}{\sqrt{1 - V^2/c^2}} = \Delta t \frac{1 - \frac{Vv}{c^2}}{\sqrt{1 - V^2/c^2}},$$

где $v = \frac{\Delta x}{\Delta t}$ — скорость сигнала в нештрихованной системе. Если $v > c$, то всегда найдутся такие относительные скорости $V < c$, для которых $1 - \frac{Vv}{c^2} < 0$, , т. е. при $\Delta t > 0$ мы получим в штрихованной системе $\Delta t' < 0$: сигнал придет раньше, чем он вышел! Требование, чтобы ни в одной инерциальной системе не могло происходить такого обращения во времени причин и следствий, и заставляет теорию относительности утверждать, что никакой процесс *сигнального* характера (позволяющий воздействовать «отсюда — туда») не может иметь скорости больше c.

Фазовая скорость может быть больше скорости света в вакууме. но она, как сказано, не может быть скоростью сигнала.

Для гравитационных волн на воде мы имели $v^2 = \frac{g}{k} = \frac{g\lambda}{2\pi}$ и $u = \frac{v}{2}$. Отсюда следует, что для очень длинных волн не только фазовая скорость v, но и групповая скорость u может быть больше c. Прежде всего это не должно нас пугать, ибо групповая скорость тоже не во всех случаях может выражать скорость сигнала. Но здесь дело еще проще: формула для v выведена без учета сжимаемости жидкости, а допущение несжимаемости тел принципиально несовместимо с теорией относительности.

Посмотрим в заключение, как обстоит дело с групповой и фазовой скоростью в квантовой механике. По де Бройлю, свободной частице, движущейся со скоростью V, соответствует ψ-волна, фазовая скорость которой $v = \frac{\omega}{k}$ определяется энергией и импульсом частицы, ибо

$$\omega = \frac{E}{\hbar} = \frac{m_0 c^2}{\hbar\sqrt{1 - V^2/c^2}}, \qquad k = \frac{p}{\hbar} = \frac{m_0 V}{\hbar\sqrt{1 - V^2/c^2}},$$

где $\hbar = \frac{h}{2\pi}$. Таким образом,

$$v = \frac{\omega}{k} = \frac{E}{p} = \frac{c^2}{V},$$

и так как $V < c$, то всегда $v > c$.

Нетрудно показать, что групповая скорость этих ψ-волн совпадает со скоростью частицы. Действительно, из выражений ω и k следует, что

$$\omega = \frac{m_0 c^2}{\hbar} \sqrt{1 + \left(\frac{\hbar k}{m_0 c}\right)^2}.$$

Но при таком законе дисперсии групповая скорость есть

$$u = \frac{\partial \omega}{\partial k} = \frac{ck}{\sqrt{1 + \left(\frac{\hbar k}{m_0 c}\right)^2}} = V.$$

{Все выводы, касающиеся распространения волновой группы, получены с помощью разложения Фурье. Почему выбран именно этот путь? Ведь волновую функцию можно разлагать по любой полной системе функций, но мы выбрали разложение именно по синусам. В данном случае ответ прост (вообще это не всегда просто). Мы сделали это потому, что мы знаем, как распространяется чистая синусоида (синусоидальная волна). Такая синусоида имеет вполне определенную скорость — фазовую скорость. Пользуясь принципом суперпозиции, мы сказали: так как нам неизвестно распространение сигнала в общем случае, то разложим его на синусоиды, для которых нам все известно. В самых лучших умах возникал вопрос: реально ли такое разложение? Относительно распространения волн в диспергирующей среде этот вопрос был поставлен не более и не менее как Стоксом и Томсоном и был освещен неверно. Правильное решение дал Релей. Рассматривая рентгеновские лучи как электромагнитные импульсы, Стокс и Томсон объясняли отсутствие дисперсии для этих лучей малой продолжительностью импульсов. Они говорили, что рентгеновские импульсы настолько коротки, что электроны не успевают раскачиваться. Уже это утверждение неверно. Но, кроме того, они делали совсем туманный вывод: следовательно, говорили они, рентгеновские импульсы нельзя разлагать на синусоиды.

Релей заметил по этому поводу следующее: можно или нельзя разлагать — это вопрос математики. Если теорема Фурье верна, то разлагать можно, а мы знаем, что эта теорема верна. Особенность рентгеновских импульсов заключается только в том, что в их разложение входят лишь очень короткие синусоидальные волны и *этим* объясняется отсутствие дисперсии. Математически оба способа рассмотрения (импульсы и синусоиды) совершенно равноправны. Это мы и можем утверждать на основании теоремы Фурье. Другое дело, когда спрашивают о физическом смысле отдельных членов ряда.

Почему в этих вопросах всегда встречаются психологические трудности? Почему их плохо понимают? Откуда берутся такие вопросы?

Есть проблемы, которые проще решать без ряда Фурье. И вот бывает так, что берут как раз такой случай, такую задачу и, запутавшись с рядом Фурье, объявляют его неверным, утверждают, что верен только простой путь, без разложения по синусам. Бывает и так, что потом строгий расчет с рядом Фурье дает другие результаты, чем простое рассуждение. Таким образом, всякий раз, когда возникает такого рода противоречия, мы имеем ошибку либо в простом рассуждении, либо в оперировании с рядом Фурье.

Примеры того, что один и тот же вопрос можно решить и более простым, и более сложным путем, часто встречаются и в других областях физики, и при этом иногда возникают противоречия. Одним из таких примеров может служить знаменитый гидростатический парадокс Паскаля. Без всякого Паскаля мы знаем, что правый сосуд тяжелее (рис. 9). Поставьте их просто на весы и убедитесь. Но закон Паскаля утверждает, что давление на дно в обоих одинаковое, а раз днища одного размера, то одинаковы и полные давления. Следовательно, заключали отсюда, закон Паскаля неверен. Это пример того, когда ошибаются в применении сложного принципа и свою ошибку приписывают недостатку этого принципа.

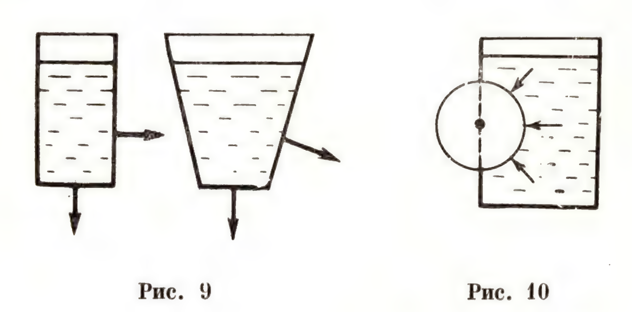


Рис. 9 Рис. 10

Обратный случай, когда несоответствие получается из-за ошибочности простого рассуждения, мы имеем во всевозможных проектах perpetuum mobile. Например, утверждается, что колесо будет вертеться, как так его правая часть, находящаяся в жидкости, «теряет в весе» (рис. 10). Раз закон сохранения энергии говорит, что вращения не будет, значит, он неверен. Здесь ошибка в простом рассуждении: забывают, что «потеря веса» есть результат давления. Для рассматриваемого «мотора» необходимо, чтобы ко-

лесо находилось сбоку, но тогда все давления проходят через ось и момент равен нулю.

В применениях ряда Фурье имеется множество примеров недоразумений обоих сортов.}[1]

ЧЕТВЕРТАЯ ЛЕКЦИЯ

(5.V 1944 г.)

Мы подробно разобрали в прошлый раз вопрос о понятии скорости в диспергирующей среде. Мы видели, что при наличии дисперсии это понятие утрачивает однозначность. Помимо фазовой скорости $v = \frac{\omega}{k}$, с которой распространяется гармоническая волна частоты ω, можно ввести понятие групповой скорости $u = \frac{d\omega}{dk}$, связанное с наличием цуга и вообще медленно меняющейся амплитуды. Применимость этого понятия групповой скорости ограничена в ряде отношений, и, в частности, оно непригодно при сильной абсорбции в среде. Однако если абсорбция достаточно мала, то им все же можно пользоваться. Кроме v и u, мы говорили еще о скорости фронта волны и скорости сигнала.

Введение всех этих понятий предполагает, что справедлив принцип суперпозиции и что гармоническая волна распространяется, не испытывая изменений. Первое означает, что уравнения, описывающие распространение волн, линейны; второе — что эти уравнения однородны. Разумеется, если k задано для всех ω, т. е. мы имеем заданный нашими уравнениями закон дисперсии, то групповая скорость может быть тотчас же вычислена. Однако это не значит, что результат такого расчета имеет определенный смысл во всех случаях без исключения. Условия, при которых u действительно характеризует распространение, мы подробно выяснили в прошлый раз.

Рассмотрим теперь еще один важный вопрос: какова скорость распространения *энергии* в среде? Пусть в среде распространяется синусоидальная волна, протяженная от $-\infty$ до $+\infty$ (рис. 11).

[1] [В фигурных скобках приведен отрывок из выступления Л. И. Мандельштама на семинаре 1936/37 г. по дисперсии и абсорбции. Это выступление непосредственно касается вопросов, связанных с групповой скоростью и изложенных в данной лекции. Записано и обработано С. М. Рытовым.]

Ее фаза (A) распространяется с фазовой скоростью. Энергия пропорциональна квадрату амплитуды, а так как амплитуда (максимальное отклонение) перемещается с фазовой скоростью, то, казалось бы, с этой же скоростью будет распространяться и энергия. Однако такое определение скорости энергии ничего не дает. Оно не удовлетворяет основным требованиям, связанным с распространением энергии. В синусоидальной волне нет *движения* энергии, она *не* переходит из точки A в точку A'. Синусоидальная бесконечная волна может быть представлена моделью, состоящей из набора одинаковых, но не связанных друг с другом маятников (Рейнольдс). В этой цепочке маятников можно создать такую последовательность фаз, что форма колебаний будет в точности соответствовать бегущей синусоидальной волне, однако никакой передачи энергии здесь не происходит. Релей заметил по поводу этой модели, что «назвать маятники Рейнольдса средой — значит сделать им комплимент». Таким образом, отождествляя скорость энергии со скоростью распространения фазы, мы не получаем главного — потока энергии из одних частей пространства в другие. В произвольном объеме, через который проходит синусоидальная волна, энергия будет оставаться все время постоянной.

Какое же определение скорости перемещения энергии следует признать более целесообразным? Рассечем волну поверхностью, перпендикулярной к направлению распространения. Если в среднем за единицу времени количество энергии справа от единичной пощадки на этой поверхности увеличивается на $\bar{S}$, а слева настолько же уменьшается (или наоборот), а средняя плотность энер-

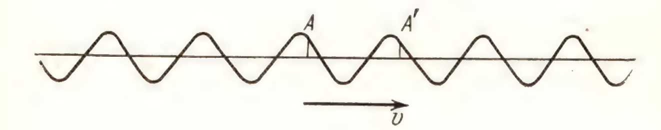


Рис. 11

гии возле рассматриваемой площадки есть $\bar{E}$, то естественно считать

$$U = \frac{\bar{S}}{\bar{E}}$$

скоростью перемещения энергии. Это определение имеет простую гидродинамическую аналогию: количество жидкости, ежесекундно протекающее через единичную площадку (ориентированную нормально, т. е. так, что количество протекающей жидкости максимально), деленное на плотность жидкости, дает скорость потока.

Однако здесь есть и существенное различие. В гидродинамике мы имеем возможность *локализовать* частицу жидкости (отметить ее краской) и поэтому можем говорить о скорости потока и в том случае, когда этот поток стационарен и однороден во всем пространстве. В электромагнитном поле такой локализации нет. Поэтому в случае неограниченной синусоидальной волны наше определение утрачивает смысл: ни о каком увеличении или уменьшении энергии по обе стороны от поверхности здесь говорить не приходится. Изменение количества электромагнитной энергии в каком-либо объеме среды можно констатировать лишь тогда, когда мы имеем *ограниченную* «синусоиду», т. е. *цуг* или *группу* волн. Тогда скорость U как раз и будет характеризовать, насколько быстро энергия проходит через поверхность.

В какой же связи находятся между собой скорость переноса энергии U и групповая скорость u?

М. А. Леонтович доказал чрезвычайно общую теорему, справедливую для любых волн. Если лагранжева функция, зависящая от величин, характеризующих состояние среды (напряженностей электромагнитного поля, смещений частиц жидкости), является квадратичной формой от этих величин и от их первых производных по координатам и времени, то в тех случаях, когда вообще имеет смысл групповая скорость u, она совпадает со скоростью энергии U, т. е. мы имеем тогда

$$\overline{S} = \overline{E}u.$$

Весьма общие предположения о существовании в виде лагранжевой функции, сделанные при доказательстве этой теоремы, все же налагают известные ограничения на свойства среды: должно отсутствовать поглощение и такие явления, как вращение плоскости поляризации. С другой стороны, для того, чтобы понятие групповой скорости имело смысл, спектр волновой группы должен быть, как мы видели, при данной дисперсии достаточно узким.

Пусть все эти условия выполнены, и, следовательно, энергия перемещается с групповой скоростью. Но мы знаем, что групповая скорость может быть отрицательной. Это означает, что группа (и энергия) движется в сторону, противоположную направлению распространения фазы волны. Возможны ли такие случаи в действительности?

В 1904 г. Лямб придумал некоторые искусственные механические модели одномерных «сред», в которых групповая скорость может быть отрицательной. Сам, он, по-видимому, не считал, что приведенные им примеры могут иметь физическое применение. Но, как оказывается, существуют и вполне реальные среды, в которых для некоторых областей частот фазовая и групповая скорости действительно направлены навстречу друг другу. Это получается в так

называемых «оптических» ветвях акустического спектра кристаллической решетки, рассмотренных М. Борном [1]. Возможность подобного явления позволяет с несколько иной точки зрения подойти и к таким, казалось бы, хорошо известным вещам, как отражение и преломление плоской волны на плоскости раздела между двумя непоглощающими средами. Протекание этого явления, при разборе которого о групповой скорости обычно вообще не упоминают, существенно зависит от ее знака.

Действительно, в чем заключается идея вывода формул Френеля?

Рассматривают плоскую синусоидальную волну, падающую под углом φ на плоскость раздела $y = 0$

$$E_{\text{пад}} = e^{i\,[\omega t - k(x \sin \varphi + y \cos \varphi)]},$$

и наряду с ней еще две волны — отраженную

$$E_{\text{отр}} = e^{i\,[\omega t - k(x \sin \varphi' - y \cos \varphi')]}$$

и преломленную

$$E_{\text{прел}} = e^{i\,[\omega t - k_1(x \sin \varphi_1 + y \cos \varphi_1)]}.$$

На плоскости $y = 0$ эти волны должны удовлетворять так называемым граничным условиям. Для упругих тел это условия непрерывности напряжений и смещений по обе стороны от границы. В электромагнитной задаче на плоскости раздела должны быть непрерывны тангенциальные составляющие напряженностей и нормальные составляющие индукций. Легко показать, что с одной только отраженной волной (или только с преломленной) этим граничным условиям удовлетворить нельзя. Наоборот, при наличии обеих волн условия всегда могут быть выполнены. Отсюда, между прочим, вовсе не следует, что должно быть только *три* волны, а не больше: граничные условия допускают наличие еще одной, четвертой волны, идущей под углом $\pi - \varphi_1$, во второй среде. Обычно молча принимают, что этой волны нет, т. е. постулируют, что во второй среде распространяется только *одна* волна.

Из граничных условий тотчас же следует закон отражения

$$\sin \varphi = \sin \varphi' \quad \text{или} \quad \varphi = \varphi'$$

и закон преломления

$$k \sin \varphi = k_1 \sin \varphi_1.$$

Однако последнее равенство удовлетворяется как при угле φ_1, так и при угле $\pi - \varphi_1$. Волна во второй среде, соответствующая φ_1,

[1] [См. том II, статья 52.]

распространяется по направлению от границы раздела (рис. 12). Волна же, соответствующая $\pi - \varphi_1$, распространяется по направлению к границе раздела (рис. 13). Считается само собой понятным, что второй волны быть не может, так как свет падает из первой среды на вторую, а значит, во второй среде *энергия* должна оттекать от границы раздела. Но причем здесь энергия? Ведь направление распространения волны определяется ее *фазовой* скоростью, энергия же перемещается с *групповой* скоростью. Здесь допускается, таким образом, логический скачок, которого не чувствуют лишь потому, что привыкли к совпадению направлений распространения энергии и фазы. Если такое совпадение имеет место, т. е. если групповая скорость положительна, то тогда все получается правильно. Если же имеем случай отрицательной групповой скорости — случай, как я уже говорил, вполне реальный,— то все меняется. Требуя по-прежнему, чтобы энергия во второй среде *оттекала* от границы раздела, мы приходим тогда к тому, что фаза должна *набегать* на эту границу и, следовательно, направление распространения преломленной волны будет составлять с нормалью угол

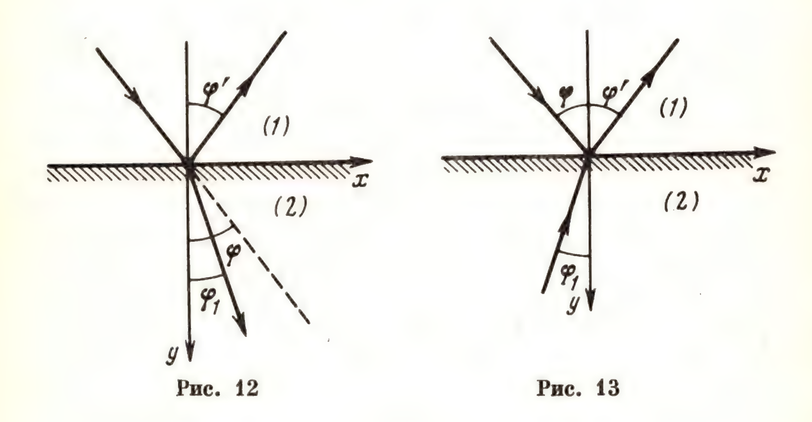


Рис. 12 Рис. 13

$\pi - \varphi_1$. Как ни непривычно такое построение, но, конечно, ничего удивительного в нем нет, ибо фазовая скорость еще ничего не говорит о направлении потока энергии.

Вопросы, которые мы разобрали, являются чрезвычайно общими — это вопросы *распространения* колебаний. Как я уже подчеркнул, они относятся к колебаниям самого разнообразного типа. По существу, я бы сказал,— это *геометрия* волнообразного движения, не связанная с той или иной физической природой объек-

та. Правда, распространение энергии уже несколько выходит из этого круга, так как это вопрос динамический.

Мы обратимся теперь к другому, тоже чрезвычайно общему вопросу геометрии (или кинематики) колебаний. Речь пойдет теперь не о распространении, а о поведении волновых процессов, так сказать, *в данном месте*. Здесь также возникают аналогичные и чрезвычайно существенные вопросы, причем опять-таки касающиеся колебаний самого различного происхождения, так что и здесь можно и нужно постараться выделить те черты, которые оказываются общими для возможно более широкого круга явлений. Прежде всего и здесь *физика* требует (и это типично) ответа на вопросы, относящиеся к *приблизительно синусоидальным* колебаниям. Показать, какого рода здесь возникают задачи, целесообразнее всего на примерах.

В любой радиопередаче, при телефонии, телеграфии, передаче изображений нам важно не синусоидальное колебание, а переносимый им *признак* передаваемого процесса — звуковых колебаний, точек и тире, сигналов бильдаппарата. Характерной здесь является именно *медленность* нарушений правильной синусоидальности: у высокочастотного колебания $y = A \cos \omega t$ происходит изменение амплитуды (или частоты) в темпе нашего разговора или телеграфных сигналов, т. е. с частотами, гораздо меньшими ω. Такие колебания называются *модулированными*. Почему целесообразно их выделить?

Прежде всего, можно отметить, что они встречаются в любой отрасли физики и техники, где мы вообще имеем дело с колебаниями. Я назову только некоторые примеры. Модуляция колебаний применяется в музыкальных инструментах (в частности, вибрато), с ней же мы сталкиваемся в явлении комбинационного рассеяния, в ультраакустике, в неравномерной морской зыби и т. д. Далее, близость модулированных колебаний к гармоническим ведет к тому, что теория их сравнительно проста и в ряде вопросов может быть развита значительно дальше, чем в случае колебаний произвольной формы. Все эти вопросы и общие положения теории модулированных колебаний были разработаны в монографии С. М. Рытова [1].

Частным случаем модуляции является так называемая *амплитудная* модуляция, при которой колебание происходит с «переменной амплитудой»

$$y = B [1 + kf(t)] \cos \omega t,$$

причем, как сказано, функция $f(t)$ меняется очень медленно по сравнению с колебанием частоты ω. В радиотелефонии изменение

[1] [Имеется в виду диссертация «Модулированные колебания и волны». Труды ФИАН, 2, 41, 1938.]

$f(t)$ происходит со звуковыми частотами, в то время как ω — высокая радиочастота. Если взять специальный случай $f(t) = \cos \Omega t$, то это означает, что $\Omega \ll \omega$. Такое колебание можно записать двояко:

$$y = B[1 + k \cos \Omega t] \cos \omega t = B\left[\cos \omega t + \frac{k}{2} \cos(\omega + \Omega) t + \frac{k}{2} \cos(\omega - \Omega) t\right].$$

Это математическое тождество, т. е. два совершенно равнозначных представления одного и того же процесса. Речь идет теперь о *действии* такого колебания на *приемник*, т. е. на устройство, представляющее собой обычно *гармонический* анализатор. Вы увидите, какую существенную роль играют во всей постановке вопроса о модулированных колебаниях свойства приемного устройства и как забвение этого обстоятельства приводило и приводит к различным высказываниям, большей частью лишенным содержания. Так, например, Флеминг утверждал, что «реально» в колебании y нет трех гармонических колебаний, а есть только *одно* колебание частоты ω, но с переменной амплитудой. Отсюда он делал вывод, что нет необходимости раздвигать радиостанции на определенные частотные интервалы. Речь шла, таким образом, о чрезвычайно существенных практических вопросах [1].

---

[1] [На этом закончилась последняя лекция Л. И. Мандельштама. В его конспектах намечены те вопросы, которые он предполагал рассмотреть в дальнейшем. В одном из таких конспектов записано следующее:

«Мы рассмотрели колебания, близкие к синусоидальным, в смысле их *распространения*. Теперь разберем их в смысле их *действия в данном месте*. Что значит «действия»? Какие здесь задачи? — Действия на приемник, анализатор. Опыты: I) Камертон a действует на настроенный камертон b; a_1 не действует на b; $a + a_1$ действует на b так же, как и a, а *на уши* совсем иначе: биения. II) a_1 не действует на b, а при прерывании действует. III) Частотомер. С какого рода приемниками мы имеем дело в физических вопросах? Эти вопросы, с одной стороны, глубоко принципиальные (белый свет), с другой — практические (дискуссия с Флемингом, частотная модуляция)».]

ОГЛАВЛЕНИЕ

Академик
Леонид Исаакович Мандельштам

**Лекции по оптике,
теории относительности и квантовой механике**

Утверждено к печати
Отделением общей физики и астрономии
Академии наук СССР

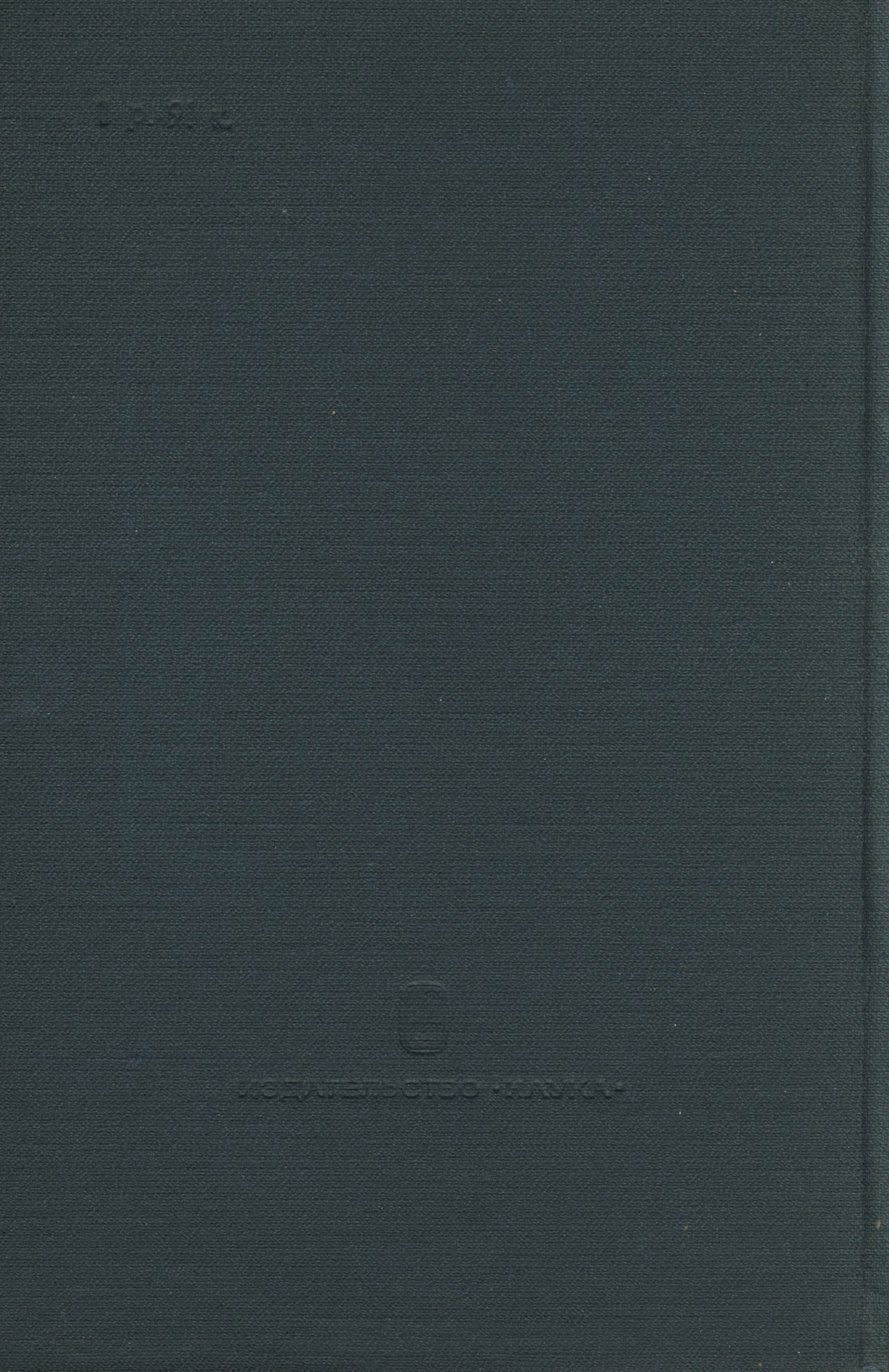
Редактор *В. П. Сироткина*
Художник *В. А. Медников*
Художественный редактор *Н. Н. Власик*
Технический редактор *Р. Г. Грузинова*

Сдано в набор 29/VI 1971 г.
Подписано к печати 12/XI 1971 г.
Формат 60×90¹/₁₆. Бумага № 2. Усл. печ. л. 27,5.
Уч.-изд. л. 24,4 Тираж 21 000. Тип. зак. 2667.

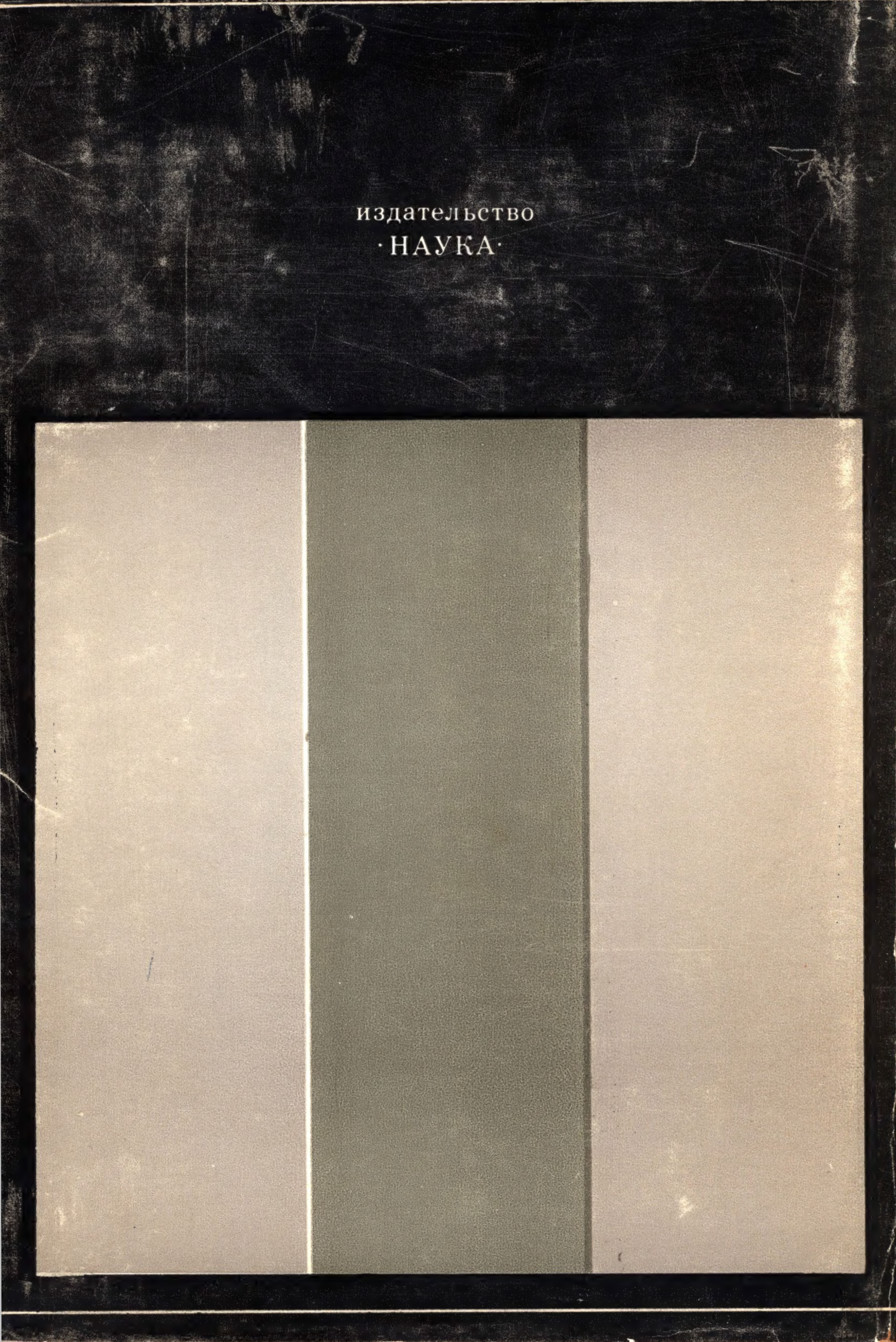
Цена 1 р. 91 к.

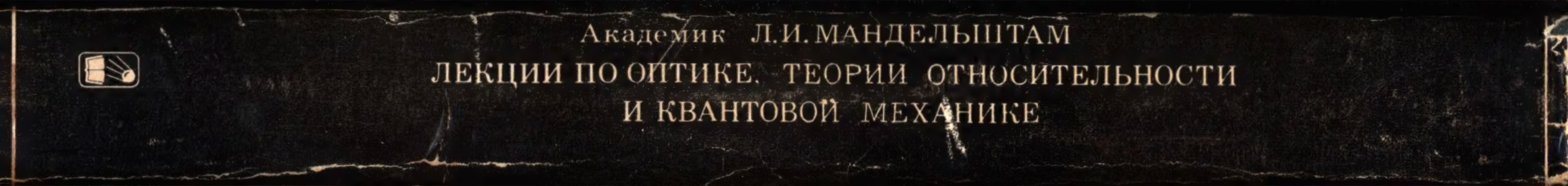
Издательство «Наука»
Москва К-62, Подсосенский пер., 21

2-я типография издательства «Наука»
Москва Г-99, Шубинский пер., 10


ИЗДАТЕЛЬСТВО «НАУКА»

1 р. 91 к.


издательство
·НАУКА·


Академик Л.И.МАНДЕЛЬШТАМ
ЛЕКЦИИ ПО ОПТИКЕ, ТЕОРИИ ОТНОСИТЕЛЬНОСТИ
И КВАНТОВОЙ МЕХАНИКЕ